AD NUMBER

AD818345

NEW LIMITATION CHANGE

TO

Approved for public release, distribution unlimited

FROM

Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational use; Jun 1967. Other requests shall be referred to Naval Postgraduate School, Monterey Ca.

AUTHORITY

USNPS ltr, 24 Nov 1971

UNITED STATES NAVAL POSTGRADUATE SCHOOL

THESIS

A METHOD FOR DETERMINING THE TEMPO OF OPERATIONS ABOARD AIRCRAFT CARRIERS THROUGH REGRESSION ANALYSES OF ACCIDENTS

by

Robert Eugene Gasser
Lieutenant, United States Navy

June 1967

A METHOD FOR DETERMINING THE TEMPO OF OPERATIONS ABOARD AIRCRAFT CARRIERS THROUGH REGRESSION ANALYSES OF ACCIDENTS

by

Robert Eugene Gasser

Lieutenant, United States Navy

B.S., Naval Academy, 1960

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL

June 1967

Signature of Author Robert Eugene Gasser

Approved by Gary A. Tuck
Thesis Advisor

Jack R Borsting
Chairman, Department of Operations Analysis

R. F. Rinehart
Academic Dean

ABSTRACT

Although variations in the tempo of flight deck operations aboard an aircraft carrier can be easily detected, no suitable method has been developed to measure this tempo. A method, based on two assumptions, which may solve this problem is presented. That the occurrence of accidents is linearly related to the many measurable events or factors comprising flight deck operations is the first assumption. The second is that tempo has a similar linear relationship with these factors. A regression analysis of accidents from previous flight deck operations is employed to find the partial correlation coefficients for each of these factors. An equation to measure tempo is then obtained using these coefficients as the weights for the various factors. Data from a U.S.S. FORRESTAL cruise are used to find these partial correlation coefficients for certain factors. In most cases these coefficients do not test significantly different from zero. As a result their value as weights is questionable. However, this does not invalidate the concepts used to develop the measure of tempo. It is hoped that future analyses, based on data from the cruises of many ships, will show this measure of tempo valid and reliable.

ERRATA SHEET

Thesis: A METHOD FOR DETERMINING THE TEMPO OF OPERATIONS ABOARD AIRCRAFT CARRIERS THROUGH REGRESSION ANALYSES OF ACCIDENTS

The following changes should be made in subject thesis:

a. page 24, line 6: change to read "approximately" vice appxoimately.

b. page 28, line 4: change to read "straight" vice stright.

c. Page 38, fifth line from bottom: change to read "curiosity" vice curiousity.

d. Page 73, fifth line from bottom: change comma to semi-colon after "exist".

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Flight deck operations aboard a modern aircraft carrier at sea vary from a peaceful lull to a pace so rapid that they border on the limitations of the machines and the men who handle them. This varying degree of pace is known as "tempo of operations." Attempts have been made to measure this tempo, however, due to tempo's complex nature these attempts have succeeded in measuring only a part of tempo. This has unjustifiably been considered a measure of the whole. Currently the measure of tempo is the total number of sorties recovered in a given period. If a single factor is to be used to measure tempo, this is perhaps the best. Nevertheless, it is not a reliable measure of the tempo of all the flight operations, because intuitively, tempo is not a single factor such as the number of sorties flown, but rather is made up of many such factors. Consideration of the circumstances under which operations are carried out must also be made. It is the purpose of this thesis to develop and examine a method to measure the tempo of flight operations aboard aircraft carriers; a method which not only considers the number of sorties flown but also the surrounding circumstances.

A measure of tempo of carrier operations which is reliable and statistically sound is desperately needed by the Naval Aviation Safety Center, Norfolk, Virginia. In the past, the Safety Center has investigated every phase of carrier operations in an effort to determine the

causes of accidents. These investigations were limited in a statistical sense because the Safety Center lacked the ability to carry out a highly sophisticated statistical study. In the near future, this will no longer be true. However, in order for the Safety Center to statistically analyze carrier operations, there exists a requirement that these operations be accurately measured. The number of sorties flown does not seem to be a good measure of carrier operations because it is incomplete. Although it does give a good indication of how the tempo of operations is varying, it is not, itself, a measure of tempo. The number of sorties flown is simply one of the factors which combine to make up tempo. Therefore, if the Safety Center bases a study on this single factor, the study cannot be complete and will be unreliable.

Carrier operations can be divided into numerous areas. Each of these areas can be considered to have a certain tempo. Investigating the tempo of all these segments of operations is a tremendous problem. As a result, this thesis is strictly limited to developing a measure of tempo of flight deck operations (hereafter usually referred to just as tempo or tempo of operations). This is the area of highest tempo and is the most important. Even though this thesis is strictly concerned with flight deck tempo, there is no reason why its concepts cannot and should not be extended to cover the tempo of operations for the entire carrier if these concepts prove to be valid.

Since the flight deck is such an important area of operation, numerous records are kept of its activities. For the statistical analysis for this thesis, many such records were obtained from U.S.S. FORRESTAL, CVA-59. These records covered FORRESTAL's

Mediterranean cruise from the date she deployed, 16 August 1965, until her return to Norfolk, 7 April 1966. Several of the records obtained were official reports. These were considered very accurate. Other records were checked by the author and corrections were made when discrepancies were discovered.

As in any *post facto* analysis, the accuracy of records is of some concern because of the possibility of human error. An even more formidable problem encountered in such an analysis is the same as that which is met in any uncontrolled experiment. The data collected is not the most suitable for the type of analysis to be carried out. The only solution to such a problem is to limit the analysis to that area which can be satisfactorily covered by the data available.

Since the statistical analysis advanced in this thesis was based on *post facto* data, rather than let the analysis suffer from the use of incorrect values, it was limited only to factors for which accurate data were available. Although this decreased the scope of operations under study, it did not seriously affect the type of analysis or the basic ideas under development.

Tempo is defined as the rate at which something is being done. However, in order to meet the needs of the Safety Center and thus to include the circumstances under which operations are carried out, the definition has to be broadened. It must include not only those factors which are done at a variable rate, but also, it must include those factors which affect a carrier's ability to function at a particular rate (e.g., included might be such factors as weather conditions, the number of

aircraft available for operations, etc.).

Tempo of operations is difficult to measure because it has an abstract nature. An observer can easily tell when there is a change in tempo, but he cannot determine the amount of the change because he has neither a unit of measurement nor a reference point. Combining this and the fact that factors used in measuring tempo are of different units, one is attracted to the idea of using a unitless quantity, the index number, to measure tempo.

An index number is defined as a statistical device for measuring changes in groups of data. The groups of data to which an index number is best applied should fluctuate widely, but should exhibit definite and measurable tendencies. Index numbers are actually relative or absolute averaging devices which serve as yardsticks of comparative measurement. Index numbers usually measure fluctuations during intervals of time. Since tempo definitely varies with time, the justification of the use of the index number is dependent only on whether the relationship of the factors to tempo has definite and measurable tendencies. This will be investigated later by means of a regression analysis.

To develop an index number for tempo, each influencing factor which can be measured must be weighted and then all such factors combined to form the index number. The equation of this relationship is introduced in the next chapter. Chapter III discusses the exact type of relationship which is used. The remaining chapters are reserved

for investigating assumptions made by applying the techniques advanced in this thesis to the data obtained from FORRESTAL's cruise.

CHAPTER II

FACTORS INFLUENCING TEMPO

It is impossible to account for and measure all the factors which combine to make up the tempo of flight operations aboard a carrier. Even if all factors are considered, some will have such a small influence that the extra effort required to include them will not be profitable. Nevertheless, a statistical study of every measurable factor must be made in order to avoid the possibility of eliminating any factor which may have some influence that is not obvious.

It is difficult to keep from counting the same factor in different ways, and yet, it is also easy to overlook important factors. In an effort to refrain from being redundant, yet at the same time attempting to be complete, the following list is presented:

1. Number of sorties flown per day
2. Number of sorties flown per night
3. Number of hours flown per day
4. Number of hours flown per night
5. Number of hours in a work day
6. Average number of sorties per pilot per day
7. Average number of aircraft launched or recovered on a cycle
8. Number of days at sea during current at sea period
9. Number of personnel aboard (pilots, deck crews, and maintenance crews)
10. Number of aircraft aboard
11. Number of different aircraft launched on a day (by side number, e.g., three different A4C's, 501, 506, and 512; six different F8E's, 801, 805, etc.)
12. Number of days on current cruise
13. Number of consecutive operating days
14. Type of aircraft launched
15. Type of flights (combat, test, low-level, intercept, tanker, bombing, reconnaissance, etc., especially for ordnance considerations)

16. Weather conditions
17. Sea state
18. Average length of time available between launches (i. e., usually the amount of time for re-spot)
19. Size of carrier

This is a list of measurable factors, where the word factor is used to denote any element that may influence tempo whether measurable or unmeasurable. After a factor is introduced into the tempo equation, it can then be referred to as a variable in the sense that it has different values for different situations.

Looking at the above list, it may be noted that both sorties and hours are included. Although they measure almost the same contribution to tempo, there are differences which may be taken into account. Also, of great importance is the fact that both day and night operations are considered. This is because tempo can definitely be observed to slow during night operations.

The number of hours in a work day is a partial measure of the effort expended by the personnel on the flight deck just as the average number of sorties per pilot per day partially measures the over-all pilot effort. Additionally, the average number of aircraft launched (or recovered) on a cycle, the number of aircraft aboard, the number of different aircraft launched on a day, the type of aircraft launched, the type of flights, and the average length of time available between the launches directly affect the tempo of the ship through the pilot or deck crew or both. Weather conditions and sea state must be considered because they are external factors which can greatly limit a carrier's operating ability.

Factors such as the number of personnel aboard, the number of aircraft aboard, and the size of the carrier are strictly inherent to a particular carrier. These factors must be considered, however, when comparing different carriers because they affect a carrier's ability to function at a tempo different from that of another carrier.

When investigating an entire cruise, additional factors such as the number of days at sea during the current at sea period, the number of days on a current cruise, and the number of consecutive operating days might also be added to the list of factors affecting tempo. Although these factors have only an indirect effect, it is possible that such factors have more of an influence on tempo through morale, than would appear evident.

Other factors were considered but were eliminated from the list because they were impossible to measure accurately. Some of these factors were training, for both pilot and deck crews, job hazards, mental capacity of an individual for his job, living conditions aboard ship, amount of pressure from the job (both mental and physical), number and type of watches and extra duties, the amount of time the ship had been at sea in the last few years, the number of days of liberty in a given period, and the type of liberty ports visited.

Assuming a relationship between tempo and the list of factors, the following equation can be written:

$$T = a_1x_1 + a_2x_2 + \cdots + a_ix_i + \cdots + a_nx_n$$

In this equation, T is the index of tempo of operations as a function of n variables, x_i is the value of each variable, and a_i is the weight of

each variable. At this point the degree of the x_i is undetermined. Not only is it possible that the x_i are of a higher degree, but they may also be formed from the combination of two or more factors.

Since an index number is unitless, each $a_i x_i$ multiplicative pair on the right side of the equation must also be unitless. This can be accomplished in one of two ways. If x_i is in hours, the a_i must have the units of a unitless weight per hour. This idea has to be extended to cover all variables. Another approach is to make the x_i quantity unitless by dividing by a minimum, average, or maximum value of the measured variable. This forces the a_i to also be a unitless quantity in keeping with the requirement of consistent unitless quantities on both sides of the equation. Although index numbers do exist which are made by adding inconsistent units, they should be avoided when possible since they are not mathematically rigorous.

To develop an index number which has meaning, it is necessary that those variables having a positive influence on the index have the common attribute that they increase in value as the degree of difficulty to carry out operations increases. Thus, if the higher weights are also given to the most influential variables, this quality of the a_i plus that of the x_i results in a higher valued index of tempo when the tempo is actually higher. There are, however, certain variables which have a negative influence on tempo, that is, as the variable increases in value, tempo actually decreases. An example of this is the amount of time available between launches. Such variables must be taken into account. This can be done by giving them a negative weight. Chapter III covers this in detail.

One problem arising when considering the variables occurs in determining the span of time over which they will be considered. It is easier to work with the normal operating day, considering 24 hours as the maximum, since most records and reports cover this period. Another problem one encounters is the fact that certain variables such as weather and sea state change, sometimes radically, in a short period of time. For these variables, an average value must be determined for the period of operations.

The reason for including many variables, such as number of hours flown, number of sorties flown, and the number of hours in a work day, is obvious. Intuitively, these variables are directly related to tempo. They are also consistent in the fact that they seem to increase as the tempo of operations increases. When measuring these variables, an average value, based on data from several carrier cruises, can be easily determined for each. Using this average value for each variable in the denominator and putting the value of the variable as it occurs on each day in the numerator, a unitless value is determined which is greater than or equal to zero and normally close to the value of one. In this way the x_i can be transferred into a unitless ratio. For example, if the computed average number of sorties for a given type of ship is 200 sorties in a 24-hour day, and a ship launches 150 sorties on a particular day, then its x_i value for sorties that day is 150/200, or 3/4.

The average value of a variable is used in the denominator to make the x_i unitless. A more useful quantity to use is a maximum value beyond which the ship will not operate. This has the desirable quality of normalizing the x_i value between zero and one. The problem with using this, however, is that it tends to affect the amount of influence a variable has on the tempo index number. This is seen by comparing two different variables. If the number of hours in a normal operating day is 12, and the maximum attainable is 24, then the average value for the x_i is 12/24, or 1/2 for this particular variable. However, for other variables, it may be different. For instance, if the average number of sorties flown per day is 100 and the maximum number possible is 400, this gives this intuitively important variable an average x_i value of 100/400, or 1/4. Since this average x_i value is less than that of the first variable considered, its influence on the tempo index number, on the average, is also less simply because of the manner in which it is measured. This is undesirable because the weighting factor alone should determine the amount of influence of each variable. By using the average rather than the maximum value, this is accomplished to a much greater degree.

The justification for including other variables in the tempo equation, although not obvious, can be easily shown. The type of aircraft launched is included in a study of tempo since the handling, launching, and landing of one type of aircraft is more difficult than another type, e.g., aircraft with nose wheel steering and nose gear launching capabilities are generally easier to handle than those without. In this case

the most difficult aircraft to handle is assigned the highest value in a scale, with the other aircraft assigned values below it accordingly. Each value is the scale number for a particular aircraft. Multiplying the number of aircraft launched by its scale number, then dividing the sum of all these values by the total number of aircraft launched times the average value of the scale number, produces the desired ratio for x_i.

Another variable easily justified is the type of flight, or the mission, of the launched aircraft. When an aircraft is launched on a combat mission with an ordnance load, there is increased stress on the pilot, but in addition, a great deal more work is also required aboard the carrier by both the ship's crew and squadron personnel in handling the ordnance, preparing the aircraft for the ordnance hop, and loading the aircraft. This is probably the most difficult type of sortie for everyone concerned. The least difficult is a sortie launched for the purpose of keeping the pilot proficient in flying his aircraft. Between these two types of missions, there are other missions of varying degrees of difficulty. These missions have to be evaluated according to degree of difficulty, with the hardest mission receiving the highest value. A time study can be made to evaluate each mission's degree of difficulty. Using the values obtained from this study and proceeding in the same manner as with the type of aircraft launched, the x_i variable can be found in this case, also.

The x_i value for weather conditions is determined by plotting visibility versus ceiling on a coordinate plane with a scale for ceiling from zero to 250 feet of altitude in 50-feet segments,

and a scale for visibility from zero to one mile in one-eighth mile segments. This scale is then marked off on each axis forming blocks. Values between one and ten are assigned to each block corresponding to the degree of difficulty in conducting flight operations. The highest value is assigned to the block where the degree of difficulty is greater, i.e., the box with (0, 0) as the coordinates of its left corner. Outside the outer limits the value is zero. Using an average value, determined from several cruises operating in different areas at different times of the year, in the denominator, any particular day's average weather condition is put in the numerator giving a suitable x_i value. To determine the average daily value of weather, a value is assigned each operating hour and the average of these is used for the daily average.

Sea state is measured by assigning a weight of ten to the worst condition, and progressively lower weights to lesser conditions. Then the procedures for weather conditions are followed. This is a separate variable because even when the weather is clear, rough seas can appreciably increase the degree of difficulty in handling the aircraft on the flight deck and in flight operations off and on the carrier.

The size of the carrier is a factor which must be included in an index of tempo of operations if the index is to be used to compare different carriers. For a particular carrier, it is a constant term in the equation. It is determined by ranking the different carriers. Since the degree of difficulty of carrying out operations increases as the size of the carrier decreases, the smallest carrier has the highest value in the ranking. This ranking is determined by personnel familiar with

operations aboard the different carriers. It can be noted that although the degree of difficulty is less for a large carrier, the larger carrier has a bigger complement of aircraft. As a result, the larger carrier may end up balanced with the smaller carrier when both degree of difficulty and aircraft complement are considered.

This chapter is primarily concerned with the factors which combine to make up tempo and the manner in which these factors can be measured. Although the basic tempo equation is introduced, the type of relationship will be discussed in the next chapter. As with any such equation, the most difficult part is weighting the factors. This will also be discussed in the next chapter.

CHAPTER III

WEIGHTING OF FACTORS

Since the only measurable parts of the tempo equation are the x_i values, additional information is needed before proceeding further if tempo is to be measured. It helps if tempo can be linked to something that can be measured. There seems to be only one element which tends to vary with tempo and to limit tempo during normal operations, while at the same time is related to the many factors being considered. This element is the occurrence of accidents. When operations are proceeding at a slow pace, or slow tempo, the likelihood of an accident is decreased. As the tempo increases, the likelihood of an accident also tends to increase. Only during very unusual operations is the tempo allowed to proceed at such a rate that it will almost surely cause a serious accident. Normally operations are curtailed when any one factor or a combination of factors increases the likelihood of an accident. This curtailment is usually in the form of a cut in the number of sorties to be flown. It can result from obvious fatigue, bad weather, rough seas, night operations, or a combination of such factors. No matter which factor is considered, it is clear that there is a definite relationship between each factor and the accident rate, and between each factor and tempo.

This thesis is written in an effort to develop a measure of tempo of operations. Since no real effort has been made in the past to do this, the only possible procedure to follow is to present a method based on certain assumptions, and then check the results to see if

the method developed is feasible. Intuitively, one can justify that tempo is a function of the factors introduced in Chapter II. Since the actual relationship between tempo and the factors is unknown, it is necessary to assume some type of relationship. The most logical relationship is of a linear form since tempo, as previously defined, should increase proportionally with the various factors at a constant rate, but not at an ever increasing rate.

Even if a linear relationship is assumed between tempo and the various factors, little can be done in determining tempo because there is no method to weight each factor. It is for this reason that the relationship between the occurrence of accidents and the factors is introduced. Obviously, as the factors increase in value, thus causing tempo to increase, the risk of an accident also increases. It is not necessary to assume a linear relationship between the occurrence of accidents and the factors, because this can be checked by least squares procedures which will be introduced in Chapter IV. If a linear relationship, or one that can be approximated by a linear relationship, does evolve after many cruises have been checked, it is possible to assume that tempo has the same type of linear relation. Of course, this is a very large assumption, but it must be remembered that there is definitely a close relationship between the occurrence of accidents and tempo, and that a measure of tempo is needed only in relative, not absolute, terms.

If the relationship between the occurrence of accidents and each factor can be approximated by a linear function, and a similar relationship is assumed for tempo, then the weighting problem can be easily

solved. First, however, it is necessary to set up a multiple linear regression model using accidents as the dependent variable and the values for the influencing factors as the independent variables. This model is based on the following equation:

$$y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_ix_i + \cdots + b_nx_n.$$

In this equation, y is the sum of accidents, x_i is the measured value of the influencing factor, b_i is the coefficient weighting each of the factors (usually called the regression coefficient), and n is the total number of influencing factors. From this model, the linear regression coefficients of each of the independent variables can be derived. Although the regression coefficients satisfy the accident equation, these coefficients cannot be used in the index of tempo equation for two reasons. First, if the regression coefficients are used with the natural x_i values, the index of tempo equation simply ends up as the accident equation. This can not serve as a measure of tempo. On the other hand, if the natural x_i values are normalized by dividing each by some pre-determined average x_i value, the equation then becomes mathematically meaningless, since the x_i quantities will be unitless and the regression coefficients will not.

A second and more important reason the regression coefficients cannot be used as weights stems from the nature of the regression coefficient. The regression coefficients dictate the amount of increase in the dependent variable resulting from a unit increase in the indicated independent variable, while the other independent variables are held constant. If two of the independent variables are highly interdependent,

then they may each have approximately the same regression coefficient. This coefficient cannot be used as a weight because it causes the underlying element, which actually causes accidents, to be added twice. An example of this can be seen from the factors of the carrier model. The number of hours flown and the number of sorties flown are highly correlated and can have appxoimately the same regression coefficient. Although a slight difference in the accident producing qualities of each of these variables may be expected, for the most part they will be the same simply because they are measuring, in different units, approximately the same thing.

The problems with the regression coefficients lead to the partial correlation coefficients. These coefficients are relative measures of the association between the dependent variable and a given independent variable, eliminating the effect of the other independent variables. Although these coefficients are always fractional values between zero and one, they are well suited for the tempo equation because they are relative values and tempo is to be measured by an index number which is also a relative value. These coefficients are unitless. It is for this reason that the values of the factors are also made unitless by dividing each factor by some pre-determined average value. Thus, the value of the index of tempo will always be unitless since the right side of the tempo equation is a sum of unitless quantities.

The development of the tempo equation using the partial correlation coefficient is based on the assumption that tempo and accidents are linearly related to the various influencing factors. Although it

is impossible to investigate the tempo relationship, since tempo has never been measured, it is not impossible to check the assumed linear relationship between the occurrence of accidents and the factors. As previously stated, this can be done by least squares techniques.

Even if the least squares curves developed for the factors are not exactly linear, it is possible that they may still be used if they can be approximated by a straight line. If this is the case, it must be remembered that any deviation from a straight line is a source of error. Since this analysis deals with actual data, it is naive to expect that a perfect straight line will result. The amount of deviation allowable, however, can be determined by a statistical test.

After collecting data points from many cruises, if a large deviation from a straight line were observed in a least squares analysis of accidents versus an important factor, this would render the assumptions of this thesis invalid. For example, consider the curves of Fig. 1. If the least squares curves for the occurrence of accidents versus a particular factor were to look like curve A, this would show that the occurrence of accidents, although increasing initially, did not continue to increase with this factor. Thus, one of the initial assumptions would have been incorrect. It would then be foolish to adhere to the assumption that there was a direct relationship between tempo and the occurrence of accidents, since this was based on the assumption of a linear relationship between the occurrence of accidents and the value of the factors. The fact that tempo increased with most factors could not be denied because of the definition of tempo. The

reason only "most factors" were included as increasing was because certain factors, such as the length of time between sorties, had a tendency to decrease in value as tempo increased. Such linear relations would have negative coefficients.

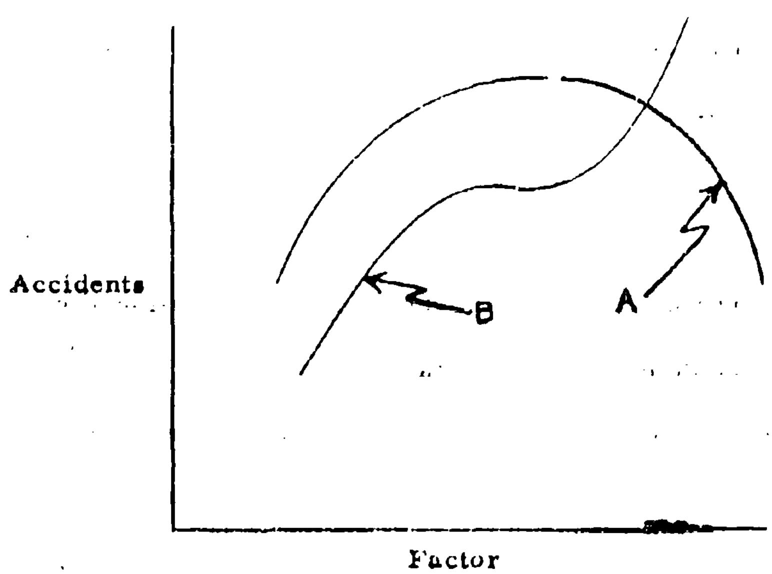


FIGURE 1

If curve A were actually observed for an important factor, the use of the partial correlation coefficients as weights would also be invalid. There would only be one value for the partial correlation coefficient, and this would be positive or negative in accordance with the predominate characteristic of the curve. If, for instance, it were positive, tempo would correctly increase in the beginning as the number of accidents increased with the quantity of the particular factor. At some point, however, tempo would continue to increase, being linear, but the number of accidents would decrease even though the factor continued to increase. Thus, in this area tempo and the occurrence of

accidents would be diverging and a great deal of error in measurement would evolve.

The situation presented by curve A should never develop if a sufficient number of data points were available to insure a reliable curve at both ends. The reason for this would be that if the upper limit of the carrier's operating capability were reached, the number of accidents occurring would be extremely large. Being an intuitive notion, this upper limit would in practice never be reached. Nevertheless, when considering factors which increase with tempo, there must be some point where the number of accidents again increased, if there previously existed a segment for which the number of accidents decreased as the factor increased. Unless this dip or level off were too great, this curve could probably be approximated by a straight line. Under such circumstances, the deviation from a straight line would indicate the amount of error which must be accepted if the assumptions of this thesis were to be adopted. An example of the type of deviation which might be expected when using actual data is shown by curve B of Fig. 1. Such a curve, with a level off or a dip, would introduce very little error and would not invalidate the assumptions of this thesis. The reason for accepting a certain amount of error in the accident curve would be due to the many outside effects on the causes of accidents which were impossible to take into consideration. Thus, it would be difficult to obtain a straight line from using actual data even if this were the true underlying relationship.

At one time it was thought that perhaps the procedures introduced in this thesis could be applied to a curve that was similar to curve A of Fig. 1. This would be accomplished by breaking the curve into three segments and approximating each segment with a stright line with only a small amount of error. However, this would be incorrect because it would contradict the assumption of a linear relationship between tempo and the occurrence of accidents for any given factor.

This chapter introduces the assumptions upon which this thesis is based. Tempo is linked to the occurrence of accidents because the two quantities have a great deal in common. The partial correlation coefficients are used as weights for the various factors in the tempo equation since they have the desired qualities. This can only be done, however, if each factor is taken to have an equation of the same degree in its least squares relationship with the occurrence of accidents. The assumption of a linear relationship is made because it intuitively seems to fit the situation for the majority of factors. This linear relationship is the basic foundation of the ideas advanced in this thesis. The only method available to check the linear relationship is to check the relationship between the measurable quantities, accidents and the various factors. This is achieved by means of least squares analyses in the next chapter.

CHAPTER IV

RELATIONSHIPS BETWEEN INDIVIDUAL FACTORS AND ACCIDENTS

To investigate the accident relationship introduced in the last chapter, all accidents related to the flight deck, including aircraft accidents and incidents, support gear accidents, and human accidents were combined daily into a dependent variable. Corresponding to this dependent variable, there was a value for each of the operating factors which was the independent variable. Using the observed data points of one factor at a time, the model

$$y_i = \sum_{k=0}^{n} a_{ik} x_i^k + \epsilon_i$$

was employed where y_i, the dependent variable, denoted the number of accidents in the i-th time period, x_i represented the value of the particular independent variable for this same time priod, ϵ_i represented the deviation of the y_i from the average number of accidents, and n was the order of the particular polynomial derived. This equation was solved by the method of least squares to determine the values of the least squares estimators, $\hat{a}_k$, for the parameters, a_{ik}. The values of the least squares estimators were then used in the model with the x_i values to determine the values of the $\hat{y}_i$. The values of the $\hat{y}_i$ were combined with their corresponding x_i value to produce points which formed the least squares curve.

By subtracting each $\hat{y}_i$ from y_i, the residual, or amount of error for each point was found. Using r_i, the value of the residual, M, the

total number of data points, and n, the order of the polynomial, the root mean square error, s, was found by means of the following equation:

$$s = \sqrt{\frac{\sum_{i=o}^{m} r_i^2}{M - n - 1}}$$

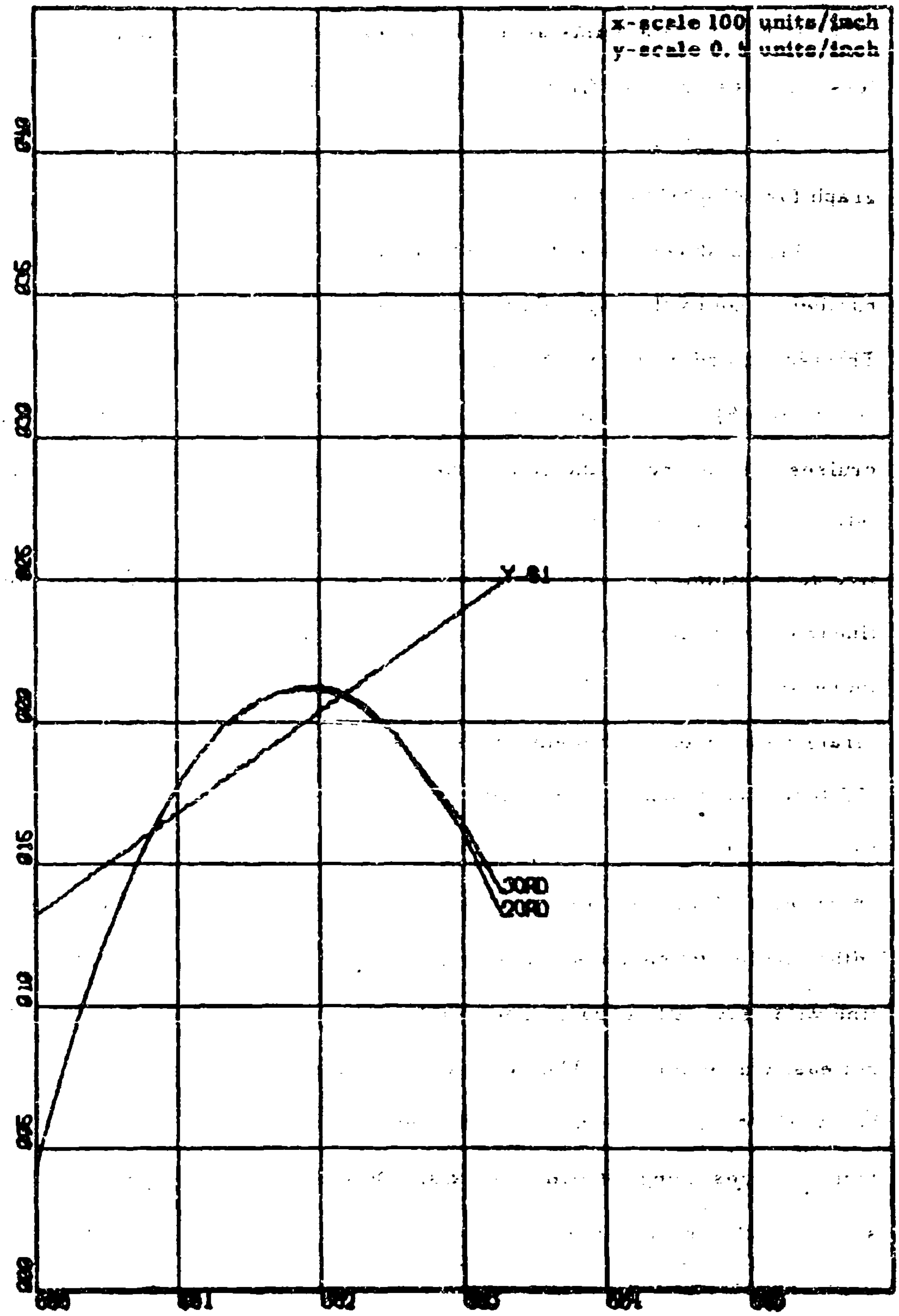
The value of s was then used to determine which order polynomial was the best fit for the data.

The above procedures were applied to certain data from FORRESTAL's last Mediterranean cruise in an effort to determine the type of relationship between the daily number of accidents and the different operating factors. Since information was not available for all factors which were previously discussed, only the following factors were investigated:

1. Number of hours flown per day
2. Number of hours flown per night
3. Number of sorties flown per day
4. Number of sorties flown per night
5. Number of aircraft aboard
6. Number of aircraft available for operations
7. Number of hours per work day
8. Average number of recoveries per cycle
9. Number of days at sea during the current operating period
10. Number of days on current cruise

All of these factors except the number of aircraft available for operations and the average number of recoveries per cycle were included in the original list. Since the data were readily available, these two factors were also investigated out of curiousity to see what their relation with accidents would be under the least squares method.

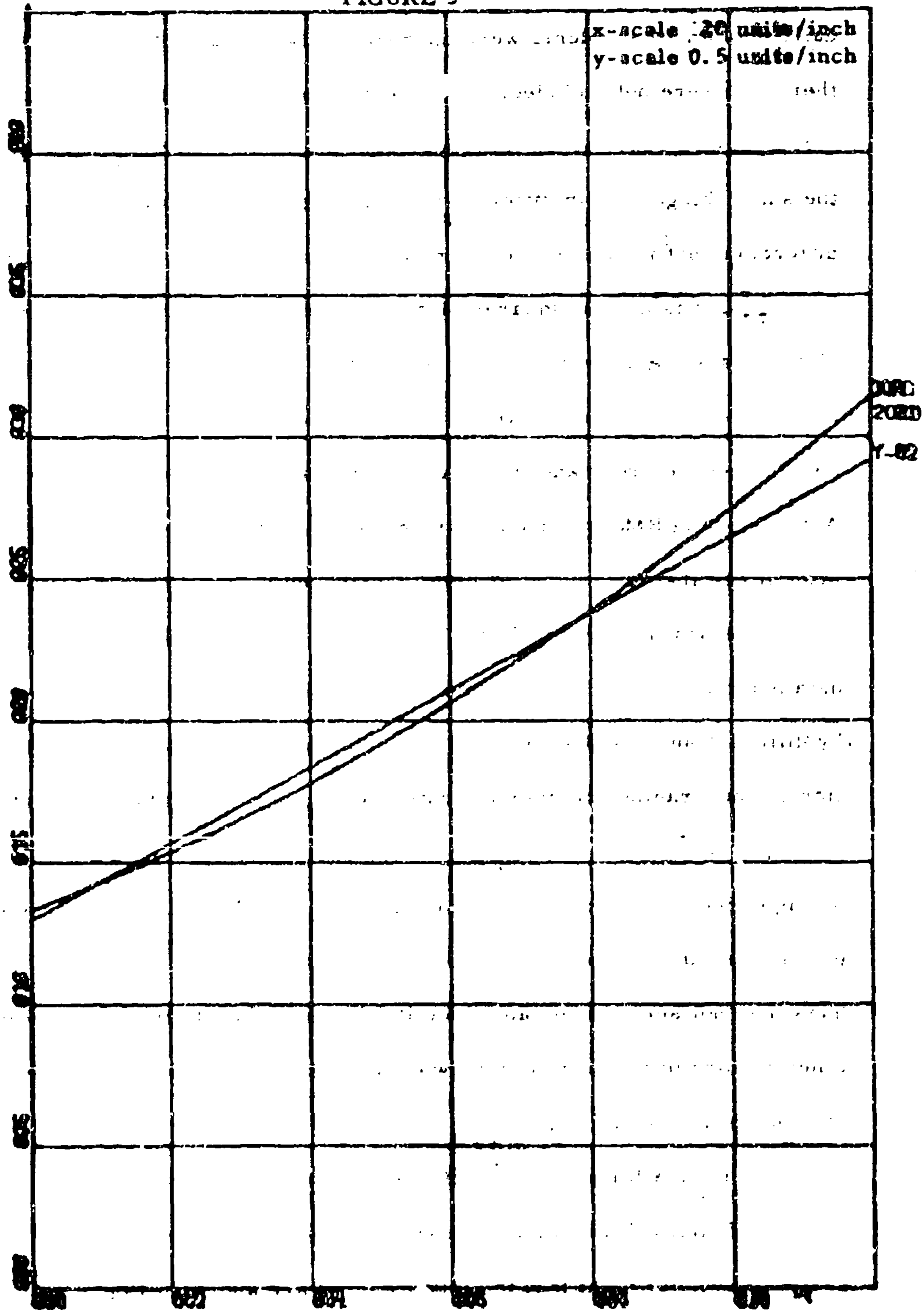
FIGURE 2



FORRESTAL operated 107 days during this cruise. Therefore, there were 107 data points available for each of the above factors. A least squares curve of first, second, and third order was fitted to the collection of data points. All three curves were plotted on the same graph for comparison purposes.

Fig. 2 shows the plot of the three least squares curves for the number of hours flown per day versus the daily number of accidents. The second order curve is the best fit according to the root mean square (RMS) error term. Although it can not be denied that after many cruises are observed, the same type curve can still result, intuitively, this does not seem likely. Beyond 240 hours the curve is based only on five data points. Evidently, these points have very little influence on the linear curve which tends to increase as the number of hours flown increases. It appears that the best curve will probably never be a straight line even after many cruises have been investigated. However, if future investigations show that intuition is correct, the curve will have a positive slope at higher values of x_i. This may occur after a short level off or even a small dip. Nevertheless, it seems that even with such inconsistencies, the curve can be approximated by a straight line with very little error. The exact amount of error acceptable is not easily determined. The situation will have to be investigated, keeping in mind that very small changes of accidents will be seen as large changes along the ordinate axis. Only then, if a curve is considered within reasonable limits, can a decision be made to use the linear approximation.

FIGURE 3

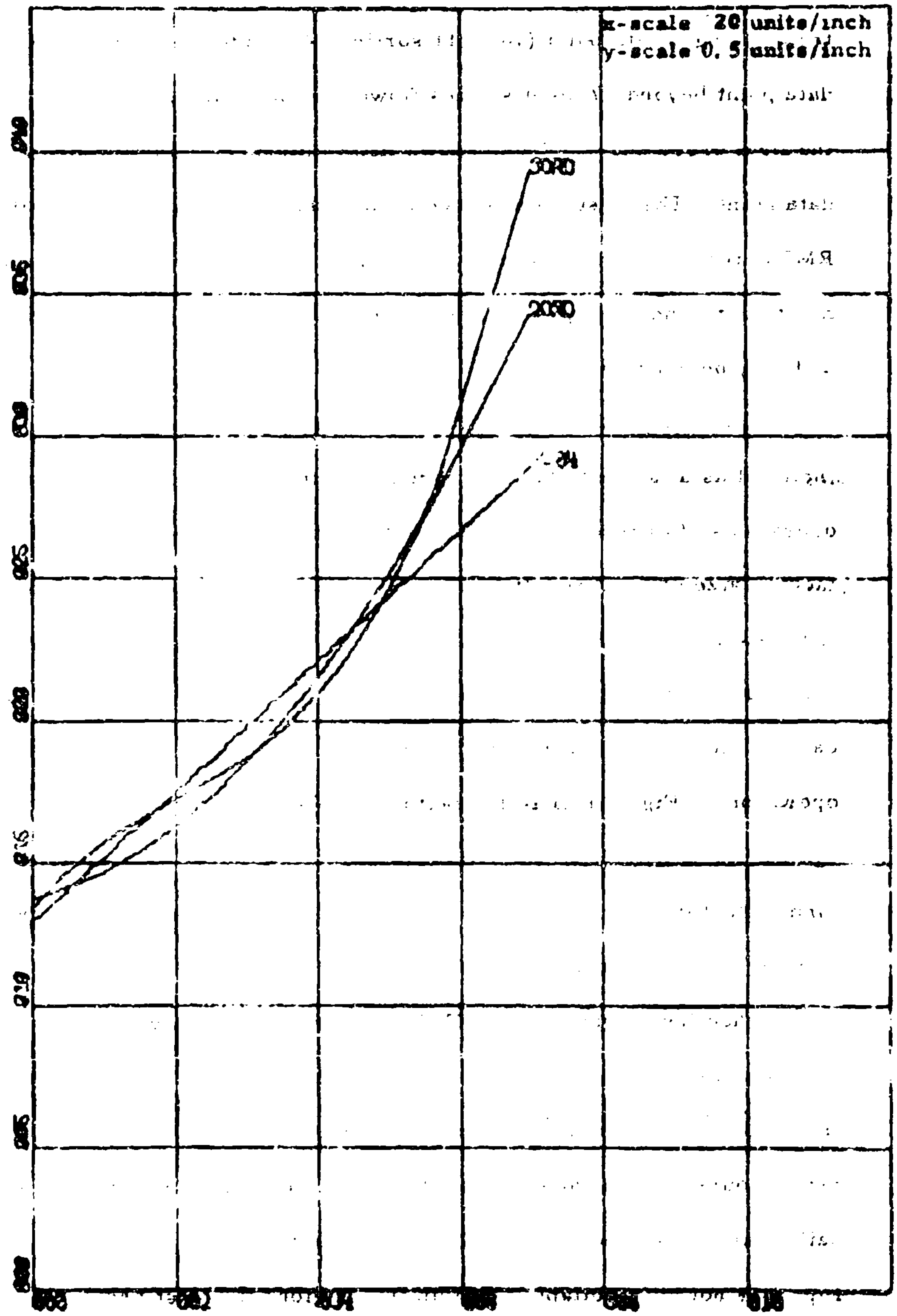


The curves for the number of sorties flown per day versus the daily number of accidents were approximately the same as Fig. 2, and therefore were not included. It was not surprising that there was not a radical difference since they were each the measure of approximately the same thing, only in different terms. There was, however, enough difference that it was decided worthwhile to include both in the research.

Fig. [illegible] is a least squares plot of the number of hours flown per night versus the daily number of accidents. The curve labeled Y-02 is the first order least squares curve. The other curve represents both the second order and third order which are the same in this case. A check of the RMS error quantity reveals that the first order is the best fit and that the second order is the second best. Although a straight line is the best fit for this factor when based on FORRESTAL's 107 data points, it is risky to assume that investigations of future cruises by different ships will show the same relationship. This first investigation is encouraging and does indicate that future investigations may at least be represented by a curve that can be closely approximated by a straight line. Even more encouraging is the fact that this curve shows an increase in accidents as the number of hours flown at night increases. This may not seem important since it follows one's intuition. However, other factors to be discussed show curves which are not so agreeable with one's intuition.

There is a large enough difference between the curves of the number of hours flown per night and the curves of the number of sorties flown per night that the latter curves are shown in Fig. 4.

FIGURE 4

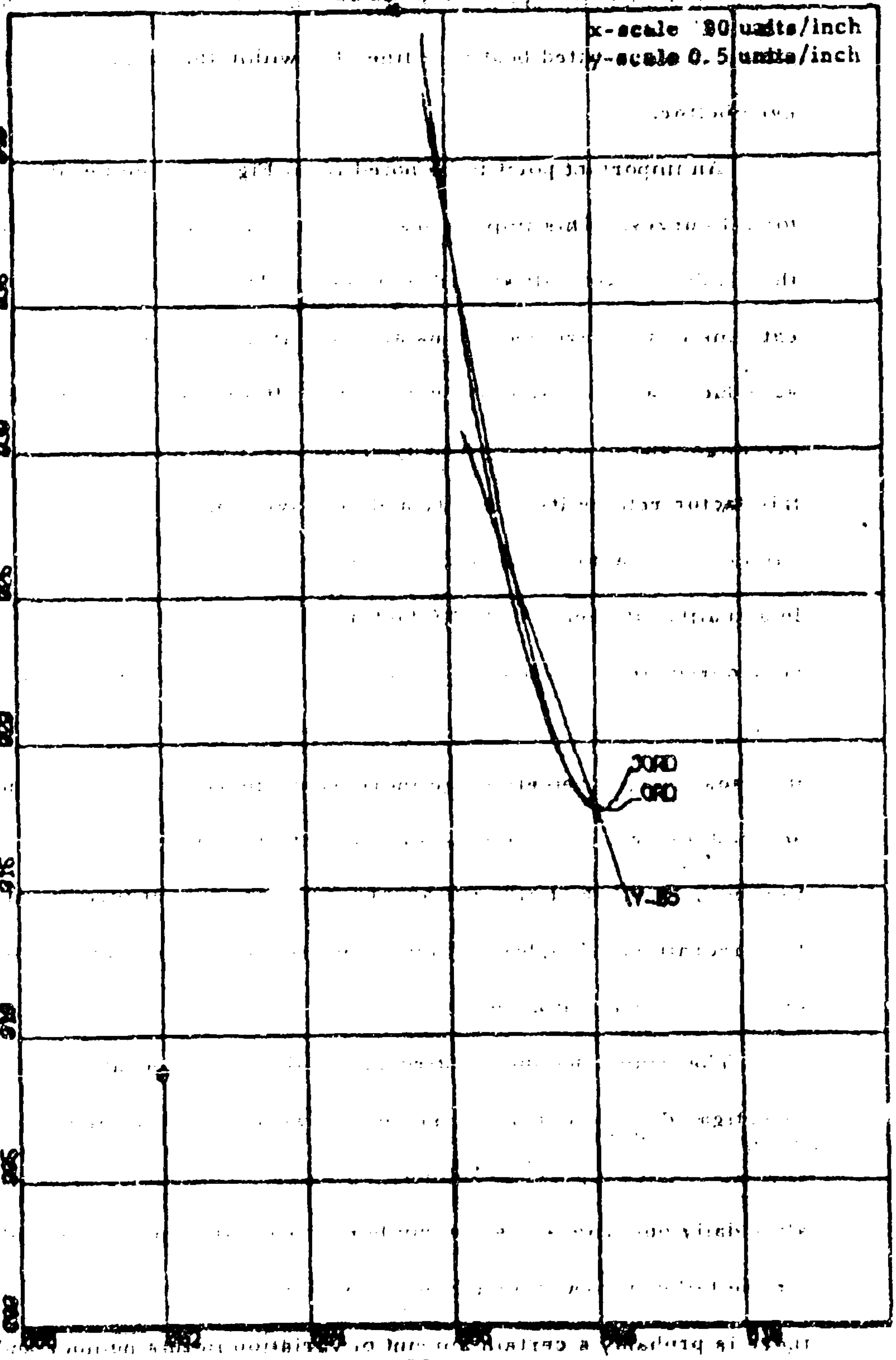


However, this difference is very unreliable because it results primarily from a single data point (70 night sorties, 4 accidents) which is the only data point beyond 57 night sorties flown for this cruise. As expected, the first order curve, Y-04, shows little influence from this single data point. The first order curve is also the best fit according to the RMS error quantity. Therefore, the apparent difference between the night hours and the night sorties' curves is actually of little consequence and may be overlooked.

In the original list of factors, the number of aircraft aboard is included as a factor to be considered only once when comparing the operations of different carriers. In this respect it is similar to the factor, size of carrier. It is possibly a balance for the comparison of a large carrier and a small carrier since operations are carried out more easily on a large carrier than a small carrier. However, a large carrier has more aircraft which tends to increase the difficulty of operations. Fig. 5 is also the factor, number of aircraft aboard, plotted against the daily number of accidents. Yet, this factor differs from the former factor of the same name, because it can be plotted on the graph since it has a daily value.

When investigating FORRESTAL's records, a daily fluctuation of the total number of aircraft aboard is noted. This daily change results from such events as visiting aircraft, aircraft diverted or sent ashore for various reasons, and aircraft lost in accidents. Since the normal daily value is between 76 and 82, the area outside this region on Fig. 5 is not very reliable. Within this region, the first order

FIGURE 5

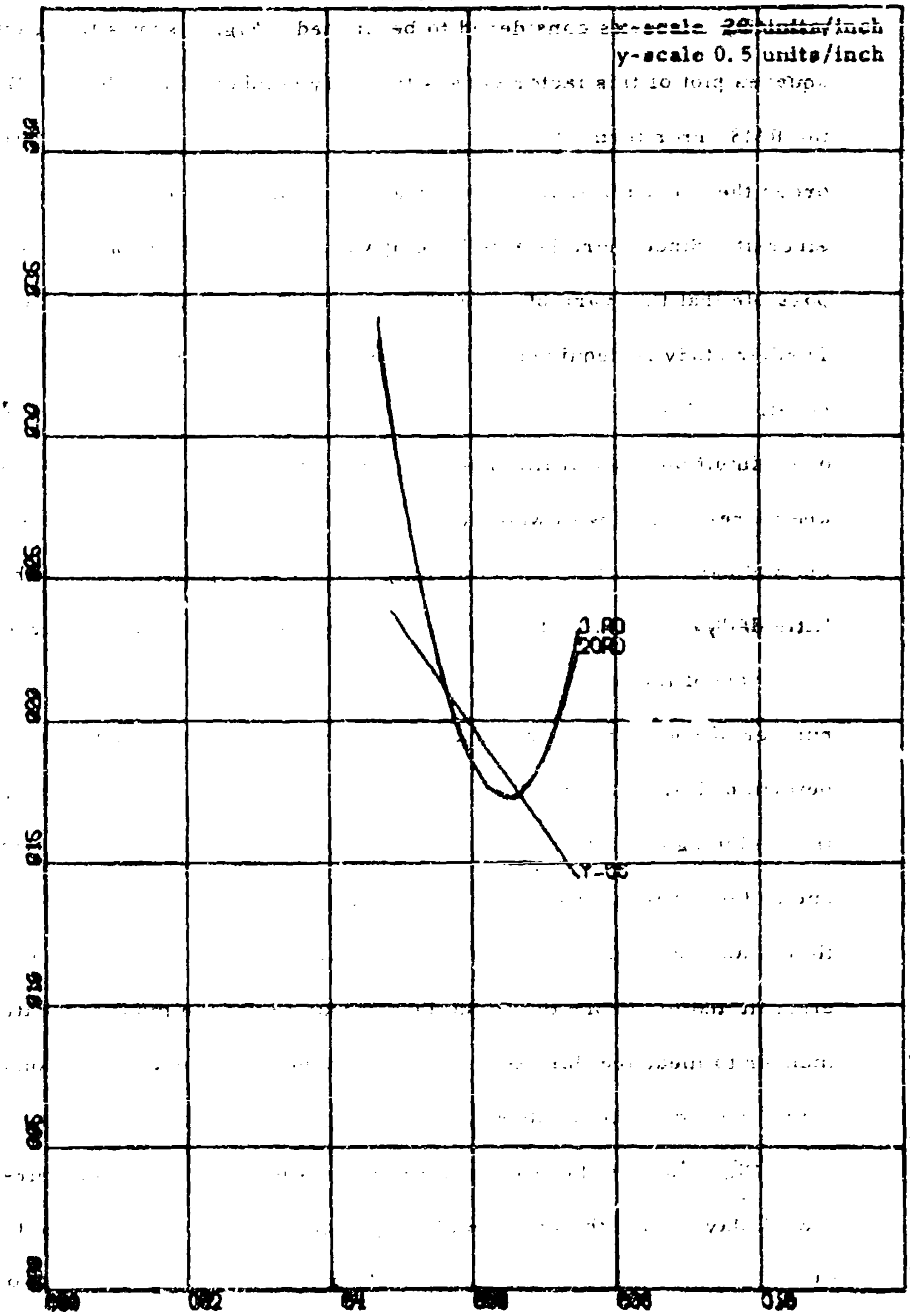


is still the best fit. As previously shown by the RMS error quantity, the entire curve is fitted best by a line, but within this region the fit is even better.

An important point to be noted from Fig. 5 is the negative slope for all curves. This implies that the number of accidents decreases as the number of aircraft aboard increases. This is certainly true in the extreme case where operations and thus accidents almost come to a stop due to an excessive number of aircraft aboard. Only through future investigations of cruises of many ships will it be possible to determine if this factor retains its linearity and negative slope. As was mentioned earlier, this factor is not included in the manner originally intended. In actuality, it seems that this factor may not be very important when considered from this point of view. It appears accidents do not necessarily increase simply because the number of aircraft increases, except in cases of an extremely large increase or decrease in the number of aircraft aboard. On a normal cruise there is very little daily fluctuation in the number of aircraft aboard. If there is a change of only one or two aircraft out of eighty or more, this seems to have very little effect on the occurrence of accidents.

The factor, number of aircraft available for operations, was investigated out of curiousity since data were available for it, and it seemed to have a possible influencing value. This factor is reported, after daily operations, as the number of aircraft expected available for the following day's operations. Since it is the expected number, there is probably a certain amount of variation in this number and the

FIGURE 6

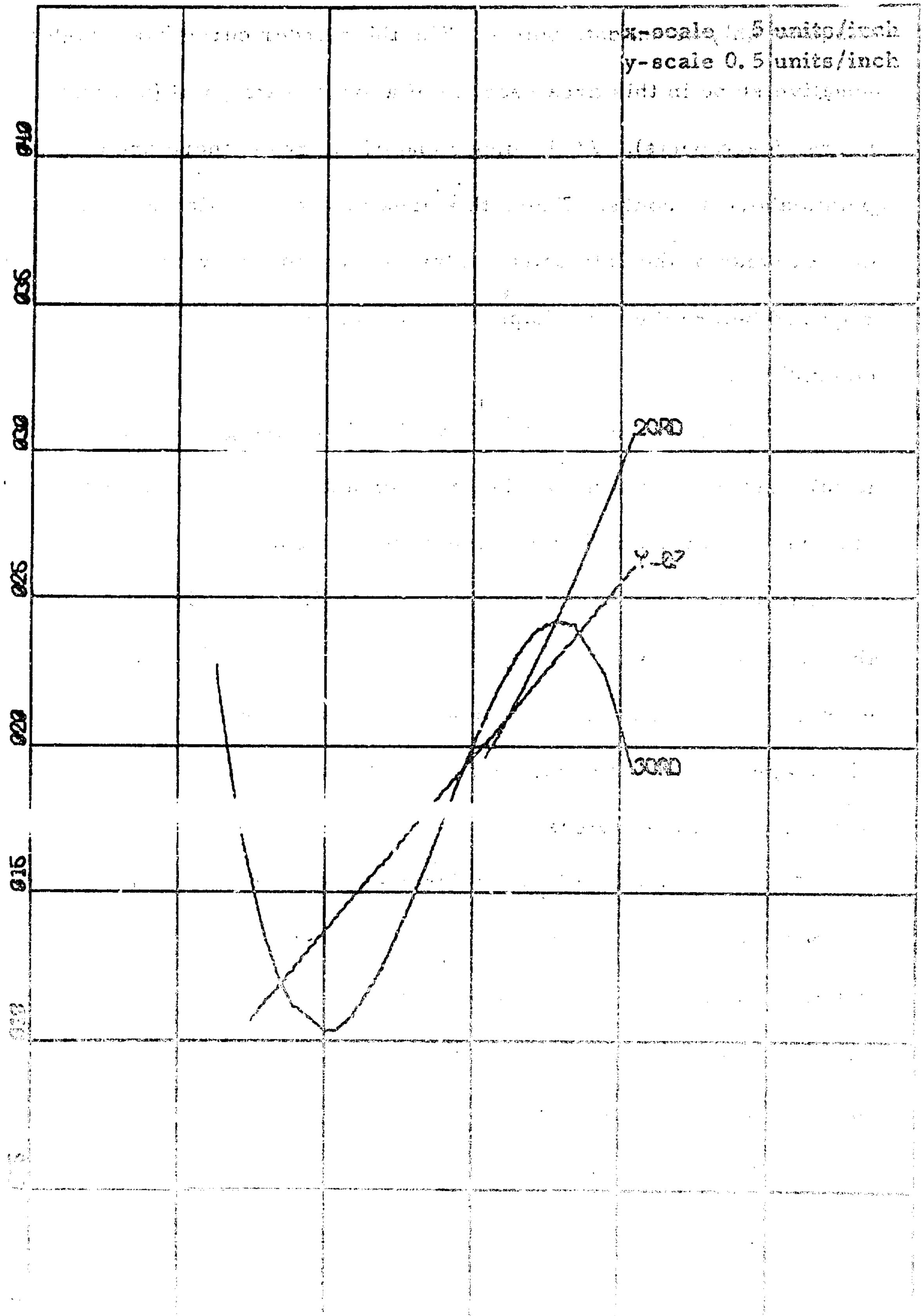


actual number of aircraft available. For this reason the value of the resulting curves is considered to be limited. Fig. 6 shows the least squares plot of this factor versus the daily number of accidents. By the RMS error term, the second order gives the best fit, and the first order the second best fit. The area of reliability is between 57 and 69 aircraft. Since there is a definite upward turn at higher values, it is possible that this variable cannot be approximated by a straight line. Further study is required if this factor is to be included in the tempo equation. Also, the negative slope, as seen in Fig. 6 does not follow one's intuition since it implies that there are less accidents when there are more aircraft with which to work. Again, as with the previous factor, the number of aircraft available may be of little significance. With very little daily fluctuation, this factor also seems to be of little importance.

One of the most important factors under consideration is the number of hours in a work day. The daily value for this factor is determined by summing the number of hours, to the nearest tenth, from the beginning of flight operations or 0800, which ever is earlier, until one and one-half hours after the last aircraft lands. This last consideration is an approximation of the time required to refuel and to respot the aircraft and to secure the flight deck. This is not the most accurate manner to measure this factor, but it is done in this manner because an actual record of the number of hours in a work day is not available.

Fig. 7 shows the least squares curves for the number of hours in a work day versus the daily number of accidents. In this case the third order is the best fit according to the RMS error term and the first order

FIGURE 7

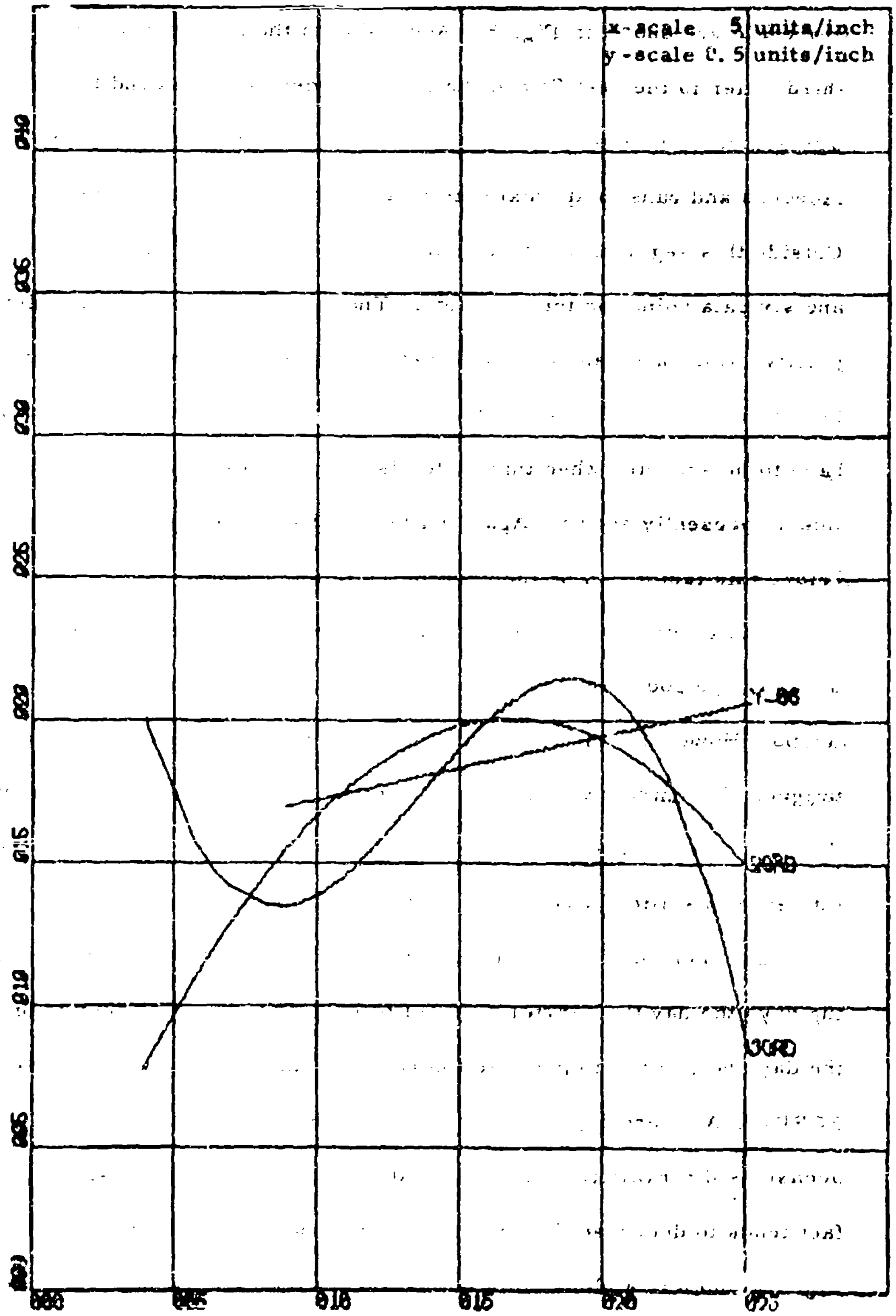


is the second best fit. Below nine hours on the graph all curves are based on only seven data points. The third order curve has a highly negative slope in this area because of a single data point (6.3 working hours, 6 accidents). At the upper end of the graph there are only four points above 18 hours. Thus, this area is very unreliable. A great deal of study of the data points in this area from many cruises will be required before the true shape of the curve on the high side can be determined.

It is difficult to ascertain how closely a straight line can be fitted to this factor. Certainly in the predominant area of operations a straight line with a highly positive slope is a good approximation. Toward the lower values there is a definite dip. Keeping in mind that these curves are based on approximations of the number of hours in a work day, which can be in error as much as one hour, again the curves of the actual data can change a great deal when obtained from the cruises of many different carriers.

The next factor to be considered is the average number of recoveries per cycle. This factor is included because it seems to have some relationship to accidents and tempo and is also easy to measure. The method of determining the average number is to divide the total number of recoveries by the total number of cycles in a day. At times it is necessary to make some modifications because certain cycles consist of only one or two aircraft. In such cases, these cycles are not counted.

FIGURE 8

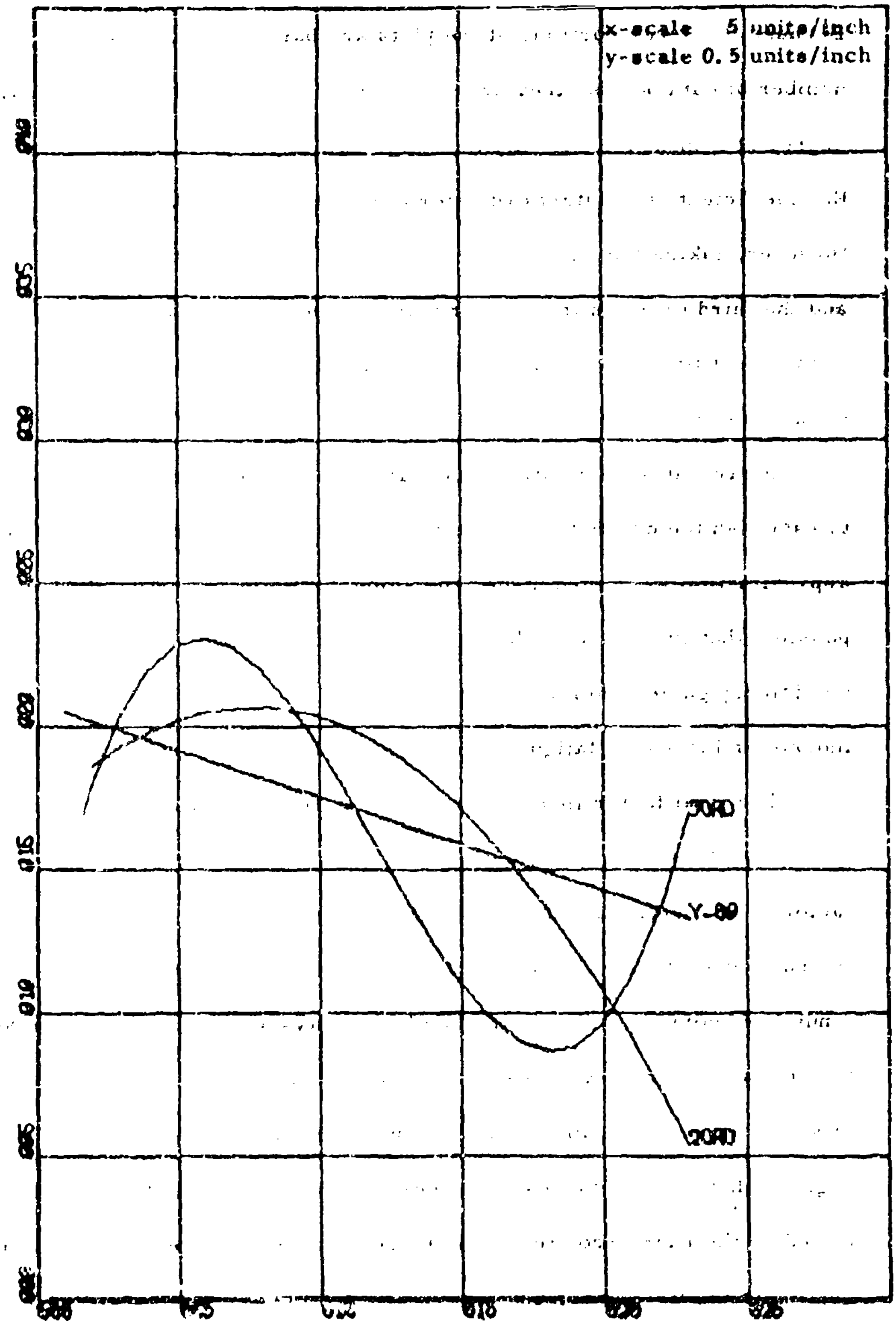


The least squares curves of the average number of recoveries per cycle are shown in Fig. 8. According to the RMS error term, the third order is the best fit and the second order is the second best fit. The segment of reliability commences around eight and two-tenths launches and runs to approximately 22 and three-tenths launches. Outside this segment there are only four data points on the low side and six data points on the high side. The resulting curve has a definite positive slope and can be approximated more efficiently with a straight line than the entire length. Of course, a certain amount of error will have to be accepted when this factor is approximated with a straight line as presently shown. Again a great deal more analysis is necessary before this factor can be included in the tempo equation.

Intuitively, the number of consecutive days at sea for a given operating period may have little direct affect on the accident rate or tempo. However, this factor may have an indirect affect on morale and fatigue. In contrast with some of the other factors, the importance of this factor is not obvious. The possibility that it may have some unknown influence is sufficient reason to justify its thorough investigation.

The method of measurement for this factor is to commence counting days the day the carrier puts out to sea, and to cease counting on the day she returns to port for liberty. Normal "at sea periods" for FORRESTAL were from four to eight days duration. On only two occasions did FORRESTAL remain at sea longer than 12 days. This fact tends to decrease the creditability beyond the 12-day point on Fig. 9, which is a plot of this factor versus the daily number of accidents

FIGURE 9

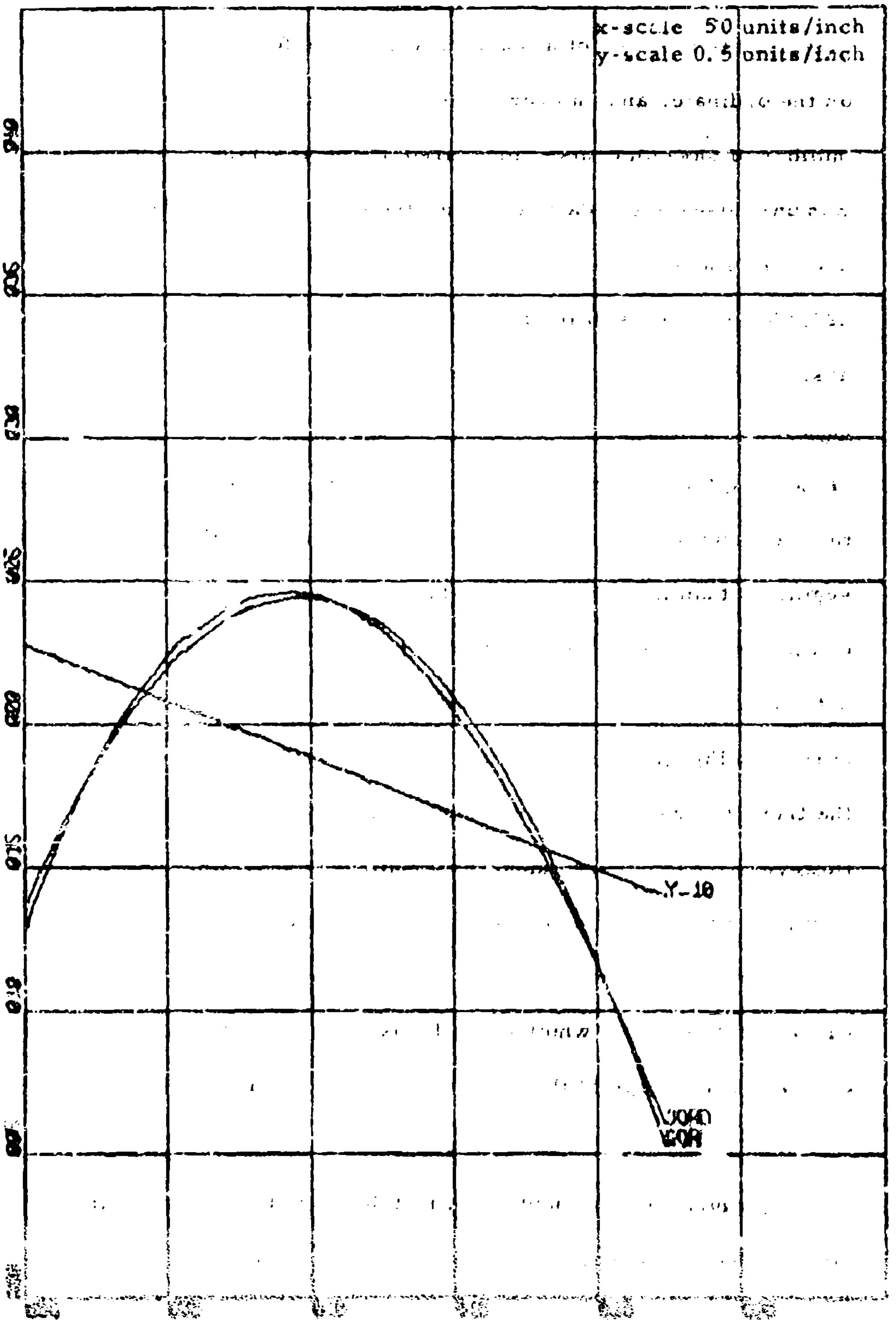


using the predicted value determined by the method of least squares. Because the upper portions of the plots are based on such a small number of data points, they are extremely difficult to analyze with any degree of confidence. If only the lower half of the plots are considered, the coefficient of the first order term may conceivably be positive. However, taking the entire curves into account, the slope is negative and the third order is the best fit with the second order having the second best fit. Thus, a linear approximation of this factor, as it stands, is unsuitable.

A great deal of additional research of many cruises would be needed to establish the correct trend for this factor at high values. If a linear approximation were then made, the slope should be obvious. It would be possible that this slope could be negative. This would indicate that the proficiency gained from a longer sea period would offset the decreased morale or increased fatigue, resulting in a reduction of accidents.

The final factor investigated with the FORRESTAL data is the number of days on the current cruise. The reason for selecting this factor is to analyze the change in the daily number of accidents with the number of days on cruise. The method of measurement is to assign a number, corresponding to the number of days away from the home port, to each succeeding day. Least squares curves are then plotted using this number and the number of accidents on each day of the cruise. Fig. 10 shows these curves. The curve which is the best fit, as determined by the RMS error term, is the second order curve. The second best fit is the first order curve.

FIGURE 10



The curves found in Fig. 10 are formed by connecting the discrete points made up of a least squares value for accidents, measured on the ordinate, and an ever increasing value corresponding to the number of days on cruise, measured on the abscissa. Since FORRESTAL was on cruise for 224 days and operated on 107 of these days, there are 107 data points with distinct abscissa values between one and 224. These curves differ from previous curves because of this distinct abscissa value. Previous curves were based on a least squares solution of data points which had as many as ten inputs for each abscissa value, and these abscissa values were relatively close together. For this reason the previous curves were much more reliable for certain segments than the curves of this factor. The use of data from many cruises will be required to develop curves which can be evaluated to determine this factor's actual relationship with accidents. As the curves in Fig. 10 are presented, they are useful as a means to observe the trend for this particular cruise. This has a great deal of value. However, these curves cannot be assumed to show what will always happen. There is probably no such curve which is so perfect as to predict this. Nonetheless, after many cruises are analyzed, several curves should evolve which could be used to estimate similar type cruises. The shape of these curves and their characteristics can only be determined by further investigation.

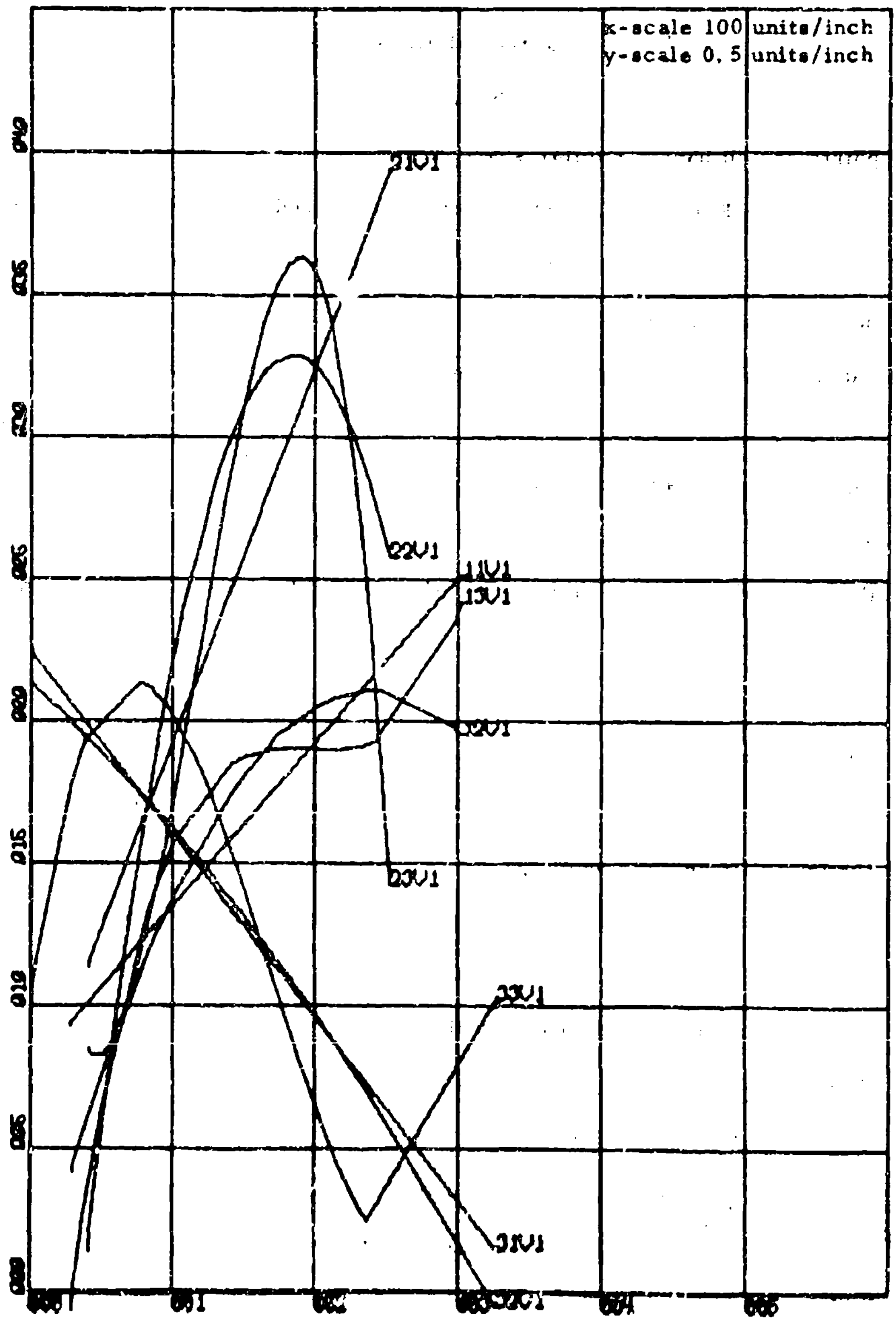
Throughout this chapter the reliability of the curves of the various factors has been questioned because certain segments of the curves

have been based on such a small number of data points. If these curves were truly representative of the relationship between the particular factor and accidents, then a sub-sample of a sufficient number of data points should have a similar shape.

Since Fig. 2 represents a factor considered important, yet shows curves of questionable validity, this factor, the number of hours flown per day, is divided into three sub-samples and each of these is investigated under least squares techniques. The first two sub-samples contain 36 data points, the last sub-sample contains 35. The curves representing the least squares investigation are shown in Fig. 11. Since there are three sub-samples, and each is checked for the first, second, and third order, there are nine curves shown. Each of these curves has a four-character label. The first number of the label designates the sub-sample where one represents the first 36 operating days and so forth. The second number of the label designates the order used to determine the particular curve. The final letter-number combination designates the factor or variable number.

The curves of Fig. 11 are in three groups. The curves of the first sub-sample are grouped in the middle of the figure. The linear curve is the best fit according to the RMS error term. It also has a definite positive slope. The curves of the second sub-sample are located in the upper area of Fig. 11. The third order curve is the best fit. Here also, the linear curve has a positive slope. The curves of the third sub-sample are located in the lower area of the figure. The

FIGURE 11


x-scale 100 units/inch
y-scale 0.5 units/inch
22V1
11V1
13V1
33V1

second order curve is the best fit. In this case, the first order curve, which differs little from the second order, has a negative slope.

The three sets of curves of Fig. 11 are completely dissimilar. Each one of the sets is also different from the set of curves of Fig. 2, however, the set of curves of Fig. 2 is the combination of the three sets of curves of Fig. 11. The difference between Fig. 2 and Fig. 11 is not surprising, and it points out the weakness of accepting Fig. 2 as singularly representative of the true relationship between the number of hours flown per day and the daily number of accidents. The curve which does represent this relationship can surely only be developed after a great deal more data are included.

This chapter shows the least squares relationship, based on data from a FORRESTAL cruise, for a selected set of factors. Fig. 11, of this chapter, is an example which shows the weakness of accepting the analysis of a single cruise as the true relationship between the factors and the daily accident rate. Since a daily value of almost any activity occurring on the flight deck can be investigated by least squares procedures, the curves displayed in this chapter are used only as indicators of possible true relationships. They also show how such analyses can be made. However, analyses of this type are based on the assumption that the data follow a trend that can be expressed by a mathematical equation. For a trend to actually exist, there should be a certain degree of association. It is therefore necessary to determine the extent of this association between the factors and the daily number of accidents to ascertain whether or not the curves have any true value. This association can be derived

by means of a regression analysis where the partial correlation coefficient is employed to give the amount of association. That the partial correlation coefficient is the same value used to weight the various factors in the tempo equation establishes two reasons for carrying out a regression analysis. The next chapter is devoted to this regression analysis, applied to the same FORRESTAL data used in this chapter.

CHAPTER V

MULTIPLE REGRESSION ANALYSIS

In Chapter IV, an effort is made to determine whether or not the least squares curves can be approximated with a straight line. In most cases it is found that the sections which are extremely curvilinear cannot be approximated with a linear curve, however, it is noted that these sections are based on very few data points. In other cases the extremely curvilinear rela onship cannot be denied. For those cases the extent of the relationship between the occurrence of accidents and the factor is questioned. Since it is possible to dovelop least squares curves between any two sets of values, whether related or not, an analysis must be carried out to determine if the questioned relationships are important. This chapter will show the results of this analysis applied to the FORRESTAL data. This is done to determine how much value to place on the inclusion of any or all of the factors in the accident equation.

To check the value of any particular relationship, it is necessary to determine the regression equation established by the trend of the least squares line. The regression equation is an average relationship. Specifically, it is the equation of the line about which the sum of the squares of the deviation of the y_i from the line itself is at a minimum. In this chapter, however, it is not the average values which are of interest, but rather the reliability of these averages. To measure the reliability it is necessary first to determine the regression equation. This whole process is accomplished by means of a regression analysis.

For a regression analysis, one measure of reliability is the standard error of estimate, s, of the regression equation. The standard error of estimate is defined as the square root of the unexplained variance of the y_i in the sample and is also known as the square root of the error mean squares.

Other quantities of interest in this chapter are the partial correlation coefficients for the various factors. Besides being needed as weights, these quantities are also abstract measures of the degree of association between the dependent variable and each independent variable. The partial correlation coefficients simply measure how well the line fits the data. However, when they are based on a sufficient number of data points so that any extraneous effect is eliminated, they give the actual influence of the factor on accidents. An additional important point to note is that the maximum value of the sum of all the partial correlation coefficients is one. Therefore, each coefficient is a percentage of the total accident influencing factors. If the total of these coefficients does not sum to one, this indicates that some of the factors, which influence the occurrence of accidents, are not included.

If the regression equation for a particular factor were best fitted by the first order with a highly positive slope, this would indicate that the occurrence of accidents increased rapidly as this factor increased. At the same time, however, this factor could have a very small partial correlation coefficient which would imply that very little of the total variance or mean squares was answered by this factor. Thus, this factor would have very little responsibility, if any, for the cause of

accidents. In such a case the regression equation would be a very poor fit for the data points and there would be a large amount of error mean squares.

In the last chapter each factor is investigated individually in an effort to determine which type of least squares curve best suits each factor. The problem with this least squares analysis is that the interaction of factors is not considered. This is indeed important but cannot be shown with graphs, so this problem is overlocked at the time and a linear curve is selected as the best over-all approximation for all of the curves. Although this linear assumption is used in the regression analysis of this chapter, the interaction of the factors will no longer be overlooked. It may seem incorrect to consider the interaction after the type of relationship is determined. This is true if the assumption and the underlying relationships are in disagreement. However, in such a case, the error will be evident from observing very small partial correlation coefficients and an extremely large error mean squares.

In carrying out an investigation which included the interaction of all factors, the factors had to be combined into an accident equation. A regression analysis was then made for the group of factors versus the daily number of accidents. This analysis was based on the regression equation for accidents which was found by use of a standard linear multiple regression computer program. Other outputs of this program of interest were: An over-all F-value based on an analysis of variance; a complete analysis of variance table; the mean, standard deviation, regression coefficient, standard error of the regression coefficient,

a computed t-value, a partial correlation coefficient, and the sum of squares added for each factor.

The F-value given for each set of factors was used in a test of hypothesis at the .05 level to determine if the combination of factors in a particular run was of any significance and thus how effective these factors were in reducing the error mean squares. The analysis of variance table showed how the total sum of squares was distributed. The mean and standard deviation were of no special use except as indicators of the situation for the factor under analysis. The regression coefficient was divided by the standard error of the regression coefficient to produce the t-value. This t-value was used in a test of hypothesis at the .05 level to see if there was any statistical reason to believe that the value of the regression coefficient given was different from zero, and thus, whether the factor under study should be included as one of the influencing factors. Another use of the regression coefficient was the determination of the slope of the relationship, between the factor and the number of accidents, from the sign the coefficient carried. The next value, the partial correlation coefficient of a particular factor, carried the same sign and could also be used to determine the slope of the relationship. This coefficient was needed, ultimately, as the weighting value for the factor in the tempo equation. It was also used in the analysis as a means of relative comparison for the factors. This comparison, however, was primarily accomplished in absolute terms by means of the sum of squares added. This quantity indicated how much

of the total sum of squares was answered by each factor when the regression analysis was carried out in the particular order indicated. The partial correlation coefficient was used to supplement this because its value was the same, independent of the order in which the analysis was carried out.

Many different computer runs were made applying the data from FORRESTAL's cruise to the linear multiple regression program. Various combinations of the ten factors, previously mentioned, were used and many different sub-samples of the 107 data points were checked. Since not all of the analyses were of equal value, some will be mentioned only briefly or collectively. The primary reason for carrying out such an extensive study was to insure that no case of significant importance was overlooked.

The first case considered was a check of all ten factors under the complete sample space of 107 data points. Although the F-value showed that the amount of variance explained was statistically significant, the analysis of variance table showed this amount to be only a little over one-sixth of the total variance. Since only eight factors were included out of 19 previously mentioned, and since there were possibly others which were not considered at that time, it was not surprising that a great deal of the varience was unexplained. It was hoped, however, that a great deal more than 17 percent would be explained since some of these factors were intuitively very important.

The problem of a large amount of unexplained variance, and thus, a large standard error of the estimate was encountered in every instance.

For this reason it seems appropriate to comment on this problem. It is impossible to account for all of the variance in a study such as this because it deals with human factors. If only 30 percent remains unexplained after all factors are included, a human factors study can be considered highly successful. This is because human situations may not lend themselves to a good statistical analysis due to their tremendous amount of variability.

With only 17 percent of the variance explained by the ten factors of this thesis, it is doubtful that as much as 50 percent would be explained even when the other factors are included. Exactly why this is the case is difficult to resolve. It is even possible that there is actually very little relation between the occurrence of accidents and the suggested factors. Thus, any mathematical relationship shown will occur only by chance. Even if some type of relationship does exist, it may be that the linear assumption is inappropriate for some or all of the factors. The possibility also exists that the sample space based on FORRESTAL's cruise is inaccurate. It can also be that this sample space is inappropriate as a single study because too much of the unexplainable human aspect entered into this particular sample space. Regardless of the actual cause of the large amount of unexplained variance, the situation can only be resolved by additional study of a larger sample space.

Returning to the first case under study, the next value investigated was the computed t-value for each factor. This value was compared against a t-value obtained from a t-table in order to test the hypothesis

that the regression coefficient was equal to zero. If the computed t-value was greater than the table t-value, the hypothesis was rejected. This was observed for only two factors, the daily number of aircraft aboard and the number of days on the current cruise. Also, these two factors had the only partial correlation coefficients which were statistically different from zero. In both cases the coefficients were negative. These results were disappointing because these two factors were intuitively of least importance. Logically, the sign of the factor and the number of aircraft aboard should have been positive. Of course, one's intuition could be wrong in either case, but it would seem that different results would be derived from a larger sample space.

The last value of interest, in the first case, was the sum of squares added for each factor. This showed that only one other factor, the number of hours flown per night, was effective in reducing the error mean squares. However, this factor had a very small positive regression coefficient, computed t-value, and partial correlation coefficient. Before the partial correlation coefficient of this factor would test significantly different from zero, a sample space of over 10,000 data points would be required, if the value of this coefficient were to remain the same with the increase in sample space. Since this would be an intuitively important factor, it would be hoped that this coefficient would greatly increase when obtained from a larger sample space, thus, requiring a great deal fewer data points to check significantly different from zero.

To further check the sample space against the ten factors, it was divided into three approximately equal sub-samples to correspond with

the beginning, middle, and end of the cruise. The first sub-sample consisted of the data from the first 36 operating days and the second and third sub-samples followed accordingly. A regression analysis was then performed on each of these sub-samples, and in addition, on the three possible combinations of these sub-samples. These six different regression analyses were carried out in order to investigate the reliability of the first analysis of the entire sample space. The results were as one might expect due to the small amount of explained variance for the total sample space. No two cases of the separate sub-samples or even combinations of the sub-samples were the same.

Only the first and third sub-samples of the three separate sub-samples proved to be significant by the F-test. In both of these cases, over 50 percent of the total variance was explained. Even with such a large amount of explained variance, only one regression coefficient and two partial correlation coefficients were significant for the first case. In the third case only one of the regression coefficients and none of the partial correlation coefficients were significant. For the second case, which was itself insignificant, all of its regression coefficients and partial correlation coefficients were insignificant.

According to the F-test at the .05 level, all three of the analyses of the combinations were determined to be significant. The amount of variance explained by each combination was slightly over one-third. In each combination only one factor was tested to have a significant regression coefficient and partial correlation coefficient. This factor was different in each case.

The concept of dividing cruises into sub-samples is important and should receive additional research. It is possible that the assumptions of this thesis are more applicable to only one phase of operations at a time. For example, the data from many cruises can be divided into two halves, one half corresponding to the first part of the cruise and the other half corresponding to the latter part of the cruise. Another possibility is to divide the data from the cruises into three phases, a beginning, a middle, and an end, similar to the sub-samples presented in this chapter. There are other possible sub-samples, but due to research limitations, only the two cases presented here will probably be worth investigating.

One problem encountered in all cases studied, which has only been briefly mentioned, was the problem of the sign of the various factors. All but two of the ten factors investigated would seem to require positive signs. In many of the regression analyses, however, this was not the case. In one case as many as seven of the factors had negative signs. This was an unusual case which was tested to be insignificant. Normally only four of the factors had negative coefficients. Two of the factors were those which seemed to require a positive sign, the number of sorties flown per day and the total number of aircraft aboard. The other two factors which were consistently negative and intuitively should be negative were the number of days at sea during the current operating period and the number of days on the current cruise.

The question of the proper sign is actually no problem until the coefficient is tested to be significantly different from zero. Until this

time the sign can be used only as a possible indicator of the direction of the trend. For this reason it is important to determine if any of the factors will become significant when other factors are deleted from the regression analysis. Deleting factors is the same as decreasing the number of constants in the regression equation. This has the same effect as increasing the sample size because it increases the number of degrees of freedom. It is entirely possible that future analyses will determine that the best equation for the accident regression requires a great deal less than the previously mentioned 19 factors. No doubt a more complex equation can be selected which fits the data more closely, but it is always best to give preference to equation types with a small number of constants if the consequence is not a markedly poorer fit. Normally the simpler the equation, the more reliable are the results.

Although reducing the number of factors can never increase the amount of variance explained, it can affect the circumstances to such an extent that the factors remaining in the analysis no longer have insignificant regression coefficients or partial correlation coefficients. For this reason two different analyses were made with a reduced number of factors. These analyses were made first using the total sample space then using the six different sub-sample spaces.

Six out of the seven cases studied, using all ten factors, were shown to be significant. Consequently, it would have been easy to make at least six out of the seven cases test significant with the reduced number of factors simply by keeping only those factors with a large sum of squares added. This would not be absolutely correct, but would work

satisfactorily for most situations. It was decided, however, that when reducing the number of factors this obvious case would be omitted. What would happen to those factors which were intuitively of greatest importance seemed to be of more interest. For this reason only those factors were retained in the regression analysis.

The first situation studied was a reduction of the original list of factors to five factors. The factors retained in the analysis were the number of hours flown per day, the number of hours flown per night, the number of sorties flown per day, the number of sorties flown per night, and the number of hours in a work day. Only two of the seven possible cases proved to be significant by the F-test. One of these cases was the regression analysis of the third sub-sample. In this case the amount of variance explained was reduced from slightly over one-half to slightly over one-third by deleting five of the factors. The regression coefficient and partial correlation coefficient of only one factor, the number of hours in a work day, tested significantly different from zero. However, this was not an improvement since this was the only factor whose regression coefficient and partial correlation coefficient also tested significantly different from zero when all ten factors were in the regression equation.

The only other case, which also tested significantly different from zero, was made up of the combination of sub-samples, one and two. In this case, the amount of variance explained was reduced, by deleting the five factors, from over one-fourth to slightly over one-fifth. None of the five remaining factors were determined to have regression

coefficients or partial correlation coefficients significantly different from zero. This was a slight degradation since one of the deleted factor's regression coefficient and partial correlation coefficient had previously been tested significantly different from zero.

After the analysis was completed, it was obvious that the situation had actually weakened when the original number of factors was reduced to only five. However, realizing that the number of hours flown per day and night was highly related to the number of sorties flown per day and night, it was decided to eliminate the former two factors. This could possibly improve the situation. Thus, only the latter two factors and the number of hours in a work day remained. These three factors were then analyzed by means of the linear multiple regression computer program. Since the elimination of the two factors which were highly correlated to the two factors remaining in the analysis had the effect of increasing the number of degrees of freedom without significantly reducing the explained variance, some improvement was observed. The improvement, however, was not of sufficient magnitude to encourage the idea that a simple equation, based on only a few factors, would be the best type of equation for the regression analysis. Only one additional case, the second sub-sample, was seen to pass from the realm of insignificance to significance.

The most important case, the one based on the entire sample space, proved to be almost significant from zero when only three factors were included in the regression equation. In this case the three partial correlation coefficients were positive and their relative size turned out

to be an ideal example of how one would intuitively weight the three factors in a tempo equation. The desired situation, after extensive data were investigated, would be to have an outcome of the partial correlation coefficients for the three factors enter at a significant level with the approximate values found in this case study. In addition at least three to five other factors should have significant partial correlation coefficients and be included in the equation, while the total sum of all the partial correlation coefficients should be around seven-tenths. If this were the case, a tempo equation using the daily normalized x_i values of each of the factors could be established. Since the values of the partial correlation coefficients were .0468 for the number of sorties flown per day, .1722 for the number of sorties flown per night, and .0187 for the number of hours in the work day, an equation used to determine the index of tempo might begin to look like the following:

$$T = .0468x_1 + .1722x_2 + .0187x_3 + \cdots$$

This would imply that night sorties were three to four times more important in determining tempo than day sorties, and that day sorties were two to three times more important in determining tempo than the number of hours in the working day. Note that since the x_i would be fractions, tempo would end up as a fraction. This would not matter because tempo should be a relative measure. The decimal place could be moved two places to the right in each weight with the consequence that tempo would evolve a whole number plus a fraction, if this were desired.

Since the linear regression analysis was for the most part insignificant, the next step was to analyze the effect of applying a second order regression equation to the entire sample space. This was performed by checking each individual factor plus its square in the regression program. The results of this effort showed that only three of the ten cases were significant. Out of the three significant cases, only the factor, number of days on current cruise, had significant regression coefficients and partial correlation coefficients. This type of analysis was used only as an indicator, just as the least squares curves were used. It was not the individual factor which was actually of interest, but rather how these factors, when combined with other factors seemed to affect the occurrence of accidents. For this reason the next procedure was to use the values of the ten factors and their squared values, for a total of 20 factors, in a single regression analysis. With the 20 factors in the equation, instead of ten, the amount of explained variance was increased one-tenth. The effect of the decrease in the number of degrees of freedom without a substantial increase in the amount of explained variance caused the entire analysis to become insignificant. This, of course, was a decline in the ability of this group of factors to effectively approximate the data. Recalling that the linear regression analysis was at least significant, surely then the relationship of some of the factors must be best represented by a straight line.

Since again the five intuitively important factors were of greatest interest, a regression analysis was made using these five factors and their squared values. This case proved to be even less significant than

the linear case with none of the coefficients checking significant. Because of this only the three factors, number of hours flown per day, number of hours flown per night, and number of hours in the work day, plus their squared values were included in the next analysis. This also checked insignificant along with all of the coefficients. The last analysis consisted of the factors, number of sorties flown per day, number of sorties flown per night, and number of hours in the work day, plus their squared values. This, too, proved to be insignificant along with all coefficients. Again this was a decline in the ability of this particular group of factors to effectively approximate the data.

Although no improvement was expected, the third order was also investigated. This was indeed the case. Actually, there was a decline in the factor's explaining abilities because only two of the ten cases of the individual factor's third order regression were significant. As expected, two of the three cases checked significant in the second order case. However, in the third order case, none of the coefficients proved to be significant.

The only other analysis carried out with the third order was to combine the first, second, and third order for each factor, a total of 30 variables, and to run this combination in a regression analysis. The results of this analysis also proved insignificant, even more so than the second order. From this it could then be said that for the ten factors selected and the FORRESTAL data, the first order gave the best over-all approximation of the data even though a great deal of variance was left unexplained, and that no improvement could be derived from using

a higher order. It should be emphasized that this was shown only for the ten factors taken together and for two different sub-samples. No assumptions could be made that all sub-samples would have the same results, although it appeared that this might be the case. Nevertheless, each sub-sample would have to be checked before any emphatic statement could be made.

The analysis, up to now, was concerned with only the independent variables or factors of the regression equation. Since one of the most important ideas introduced in this thesis was that of summing all types of accidents occurring each day for a value to be used as a dependent variable, this procedure was also checked. The types of accidents investigated were divided into four classes: accidents inflicting damage to personnel; accidents to the aircraft and associated equipment on deck, called ground accidents and crunches; aircraft incidents, which are potential accident situations which should lead to an accident but didn't; and aircraft accidents, which are accidents causing damage or destruction of an aircraft normally while in, approaching, or returning from the flying state. All previous analyses were based on the daily sum of these four types of accidents without a great deal of success. Consequently, it was decided to carry out regression analyses on each individual type of accident and all possible combinations of the four types to see if this would give better results. These fourteen different regression analyses were made using all ten factors in each case.

The most numerous accidents of any one type were the accidents to personnel. The data for these accidents were obtained from a monthly

report of injury to flight deck personnel. Every accident was counted, no matter how serious, if it occurred while a person's presence was required on the flight deck. The theory behind counting all accidents equally was that each accident was a human error, the seriousness of which was normally determined by the circumstances rather than the error itself. This basic idea also justified the adding of the different types of accidents.

Since the total number of accidents was made up primarily of the personnel accidents, very little change in the regression analysis, when using only personnel accidents, was expected. This was indeed the case with this analysis showing very little improvement in explaining the variance. Also, two coefficients checked significantly different from zero, while only one had previously done so when the regression analysis was based on the total sum of accidents. However, these improvements were inconsequential. Nevertheless, this did present an idea which might be worth investigating when a larger sample space would be available. It could turn out that better results would be obtained by using only personnel accidents in the regression equation since this might be a better indicator of tempo.

The second most numerous type of accident was the ground accident or crunch. The regression analysis using only these accidents proved to be insignificant. Also, none of the coefficients tested significantly different from zero. This was not the case for aircraft incidents. Although there were very few of these reported, the regression using only these accidents did prove significant with approximately one-fourth

of the variance explained. Since only one of the coefficients tested significantly different from zero, the regression analysis using only these accidents was not an improvement over the regression using all the accidents. The last case, a regression analysis involving only aircraft accidents, also was not an improvement over the initial case. It proved to be insignificant with none of the coefficients testing significantly different from zero.

The regression analyses using each of the different types of accidents were performed to determine if a noticeable improvement in reducing the unexplained variance could be effected. This did not eventuate, so further analyses were made using all possible combinations of the four types of accidents as the dependent variable. Out of these ten additional cases only five proved to be significant. These five consisted of the following combinations: personnel accidents and aircraft incidents; personnel accidents and aircraft accidents; aircraft incidents and aircraft accidents; personnel accidents, ground accidents, and aircraft accidents; personnel accidents, aircraft incidents, and aircraft accidents. None of these cases showed any noticeable improvement in the amount of explained variance, although two of the five cases did have as many as four regression coefficients and partial correlation coefficients which tested significantly different from zero.

By increasing the sample space the value of the coefficients can eventually be tested significantly different from zero simply because of the nature of the t-test. This, therefore, is a problem which can be rectified. However, getting as much as 70 percent of the variance

explained is something that cannot be controlled. This is the actual problem which must be faced.

The different regression analyses of this chapter were performed in an effort to produce the maximum amount of variance explained, and at the same time, use as few factors as possible. Since most analyses proved to be insignificant and none of the analyses was successful in explaining a large proportion of the variance, the desired results were not obtained and the only comment that can be made is that additional study is required. Other types of analyses could be made and some, in fact, were. A check was made to determine if there was a significant amount of autocorrelation in the series comprised of the daily number of all types of accidents. This was investigated because a large amount of autocorrelation would indicate that the number of accidents occurring each day was largely affected by the number of accidents occurring on some previous day. However, this was not the case because these autocorrelation coefficients proved to be insignificant for intervals as large as three days. With larger intervals than this, intuitively, there should be no possible association. Another analysis which was performed on the entire sample space, using the ten factors, was a stepwise regression analysis. This was not discussed in detail because the results obtained were not noticeably different from those observed when using the multiple regression analysis. However, the stepwise regression analysis should play an important part in future studies since it would be very useful in determining which factors were the most important in

explaining the variance when there was a large number of factors to be investigated.

This chapter is primarily concerned with presenting the different regression analyses which are applied to the FORRESTAL data. In each analysis, the amount of variance explained is determined to be too small. Because of this, no definite conclusion can be made using these analyses, and a need for additional analyses based on a much larger sample space is demonstrated. Although the results of these analyses are disappointing, they can be used to indicate what procedures may be adapted in future analyses. If future analyses do affirm the ideas introduced in this thesis, tempo can then be measured by an equation which is similar to the example presented in this chapter.

CHAPTER VI

CONCLUSIONS

The research for this thesis was carried out in an effort to develop a structurally sound equation to measure the tempo of flight deck operations aboard a carrier. It was determined, however, that to develop a measure of tempo it was necessary to first perform a regression analysis of accidents. It should be emphasized that the primary concern of the regression analysis was not merely to fit the historical data, with minimum residual variation, with whatever variable best did the job, but rather to identify and specify the underlying accident generating mechanism. Unless this was done, the values of the partial correlation coefficients would be useless. Thus, the final tempo equation, based on the partial correlation coefficients, would also be invalid.

Since the regression analysis of the FORRESTAL data proves to be insignificant in most cases, it is evident that the underlying accident generating mechanism is not determined from this sample space. There are several possible reasons for this. First, a regression analysis technique is mathematical and is based on the assumption that the data follows a trend that can be expressed by a mathematical equation. It is possible that a mathematical relationship does not exist, nevertheless, a larger sample space should be investigated before such a notion is accepted. A second possibility is that the linear regression analysis is inappropriate because some or all of the factors investigated do not have a linear relationship with accident occurrence. This possibility also

gives insignificant results and can only be checked by investigating a larger sample space. Another possibility is that there are errors in the FORRESTAL data that cause insignificant results. It is also possible that some important accident influencing factors are not included or that errors are made in the manner of converting the recorded data to the form required for the study. Lastly, it is possible that the FORRESTAL data contains a great deal of spuriousness.

No matter what the reason for the insignificant results obtained from the FORRESTAL data, there does seem to be justification for additional analysis to determine if the factor-accident relationship is linear. If future research is conducted, there are two possible results of importance. The factor-accident relationship can test significant using a linear regression. This implies that a linear or an approximately linear relationship exists. If so, one of the basic assumptions of this thesis is proved, thus, leading to a better measure of tempo. The other possible result from extended research of a larger sample space can be that the factor-accident relationship will test insignificant. This can be caused by any one or a combination of the previously mentioned possible causes that rendered the FORRESTAL data insignificant. No matter what the reason, however, a basic assumption is invalidated, and thus, this thesis cannot be used to provide a better measure of tempo.

The concepts of this thesis should not be rejected unless proven invalid by additional research. The analysis of the FORRESTAL data did not show the assumptions to be incorrect, but for the most part, it

showed only that nothing but insignificant results could be obtained from such a small sample space. The reason for the research upon which this thesis was based and the reason this thesis was directed toward the idea of developing a measure of tempo of operations aboard an aircraft carrier were based on a request from the Naval Aviation Safety Center. From the conception of this thesis, the possibility of invalid basic assumptions was realized and the analysis was conducted to determine if this were the case. The members of the Special Study Group at the Safety Center felt that even if an idea were developed and proved in error, at least that would be one less path that they would be required to investigate in their quest for a good measure of tempo of carrier operations.

At the initial undertaking of the study for this thesis, it was hoped that the ideas advanced would be substantiated by the analysis of the FORRESTAL data. Once it was determined that this could not be accomplished with this sample space, the only recourse was to explain in detail exactly what procedures were followed in the analysis. Thus, the Safety Center might be able to conduct similar research with a larger and more extensive sample space, if they so desired.

If the procedures advanced in this thesis are to be used for any future analyses, one important modification is suggested. If some control can be implemented in the collection of data, it is highly desirable to separate all activities in the records according to day or night. For example, currently it is impossible to distinguish from the records whether the accidents to individuals occur during the day or night. Since these accidents constitute a large percentage of the total number of

accidents in the FORRESTAL data, it was necessary to combine all accidents in a single regression analysis of day and night activities. For more reliable results, however, the better method to use in performing the analysis is to separate all day and night activities and carry out a regression analysis of each.

In review it should be noted that several parts of this thesis are based on intuitive ideas. If these ideas prove correct, a great deal of improvement will be made in the ability to measure tempo. That this possibility exists is ample justification for further study. Even if the intuitive ideas prove incorrect, additional use can be made of the different studies suggested by this thesis. The least squares curves can be used as indicators of what will happen on future cruises with regard to accidents and the various factors. The regression analyses are actually necessary for statistical accident studies even without attempting to use them in measuring tempo. These analyses will probably be conducted by the Safety Center because of their possible value in determining the causes of accidents and of predicting accidents in the future.

As a final point, it should be remembered that the index number proposed in this thesis is to be used as a single measure of relative tempo and not as a purely objective measure of tempo. These concepts are presented in an effort to obtain a better measure of tempo. They are by no means perfect and require a great deal more research. Eventually, however, they should lead the Naval Aviation Safety Center to that which it needs so desperately: a reliable, valid, and statistically sound measure of tempo of operations aboard an aircraft carrier.

BIBLIOGRAPHY

Anderson, T. W. An Introduction to Multivariate Statistical Analysis. New York: John Wiley and Sons, 1958.

Arkin, Herbert and R. R. Colton. Statistical Methods. 4th ed. College Outline Series. Paperback Edition. New York: Barnes and Noble, Inc., 1939.

Bendat, Julius S. and A. G. Piersol. Measurement and Analysis of Random Data. New York: John Wiley and Sons, Inc., 1966.

Breaux, Harold J. et al. "Stepwise Multiple Regression Statistical Theory and Computer Program Description," Ballistic Research Laboratories, BRL R 1330, July, 1966.

Croxton, Frederick E. and D. J. Cowden. Applied General Statistics. New York: Prentice-Hall, Inc., 1939.

Ezekiel, Mordecai and K. A. Fox. Methods of Correlation and Regression Analysis. 3rd ed. New York: John Wiley and Sons, Inc., 1959.

Graybill, Franklin A. An Introduction to Linear Statistical Models. New York: McGraw-Hill Book Company, Inc., 1961.

Health Sciences Computing Facility. "Multiple Regression with Case Combinations," Biomedical Computer Programs. Los Angeles, California: University of California Press, 1964.

_______. "Stepwise Regression," Biomedical Computer Programs. Los Angeles, California: University of California Press, 1964.

Large, J. P. (ed.). "Concepts and Procedures of Cost Analysis," The RAND Corporation, RM-3589-PR, June, 1963.

Modd, A. M. and F. A. Graybill. Introduction to the Theory of Statistics. 2nd ed. New York: McGraw-Hill Book Company, Inc., 1963.

Ostle, Bernard. Statistics in Research. 2nd ed. Ames, Iowa: The Iowa State University Press, 1963.

Yamane, Taro. Statistics, An Introduction Analysis. New York: Harper and Row, Publishers, Inc., 1964.